2005-05-31

Genetic differentiation in Alewife populations using microsatellite loci

Sunita R. Chilakamarri
Worcester Polytechnic Institute

GENETIC DIFFERENTIATION IN ALEWIFE POPULATIONS USING MICROSATELLITE LOCI

A Thesis

Submitted to the faculty of

WORCERSTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Master of Science in

Biology and Biotechnology

By

_____

Sunita R Chilakamarri

May 31, 2005

APPROVED:

Dr. Jill Rulfs, Major advisor _____

Dr. Lauren Mathews, Thesis committee _____

Dr. Jeffrey Tyler, Thesis committee _____

ABSTRACT

Local genetic adaptation and homing behavior in anadromous fish favors the formation of local populations across their geographic range of distribution. Spawning- and natal-site fidelity repeated over generations restricts gene flow and allows genetic differences to accumulate and may result in reproductive isolation. This could lead to progressive genetic differentiation and population structuring among different river populations.

Alewife, *Alosa pseudoharengus*, are anadromous fish which are estimated to have high rates of reproductive fidelity and hence might show population structuring among different breeding streams. Alewife are fish of economic importance since they have both commercial and recreational value. Alewife populations have been declining over the past decades and conservation measures to restore the populations have been implemented. Since maintaining genetic integrity of natural populations is one of the main concerns, identification of population structure can assist in designing appropriate restocking programs.

In this study, I used microsatellite markers developed for shad to study population structuring in alewife. Samples were collected from two sites in Connecticut and one in Lake Michigan and genetic differentiation among these populations was estimated using five microsatellite loci. My studies indicate that microsatellite loci developed for shad can be used for alewife. Results from this preliminary study indicated subtle but significant genetic differentiation among populations. This suggests that care should be taken when restocking alewife from different sites in order to maintain genetic diversity among these populations.

ACKNOWLEDGEMENTS

I take this opportunity to thank each and every individual who has contributed to my thesis work. Firstly, I would like to thank my advisors Dr Jill Rulfs, Dr Jeff Tyler and Dr Lauren Mathews for helping me complete this project. I am grateful to Dr Jeff Tyler for letting me take up this project and providing me with adequate information to go forward with this work. I am highly indebted to Dr Jill Rulfs for her patience and invaluable suggestions throughout. She was not only a mentor but also a friend who was always positive and encouraged me to be optimistic even during the toughest times. I sincerely thank Dr Lauren Mathews for her thoughtful contributions, suggestions and incredible knowledge in the field of population genetics. I also thank her for giving me access to use her lab equipment.

Apart from my advisors, I am fortunate to have the support of many other staff, faculty members and students in the department of Biology and Biotechnology. I thank Dr Mike Buckholt, Dr JoAnn Whitefleet-Smith and Dr Dave Adams for their technical expertise to troubleshoot problems. Likewise, I would like to recognize Jes Caron's help with the maintenance of the lab. I am grateful to Carol Butler who made life easier at WPI. I could always rely on her for any kind of help. Eileen Dagostino is yet another person whose assistance is worth mentioning. Special thanks to my fellow students, Elisa Ferreira, Scott Pherson and Heather Hinds for their contributions in various aspects of this project.

Thanks to the WPI Library for the gamut of information and to the Hospital for Sick Children, Toronto, Canada for the technical assistance.

Last but not the least; I am thankful to my family members for their support and encouragement all along. I am especially thankful to my husband for the cooperation and moral support without which I could not have achieved my goals.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

A review of the ecology, genetics and management of fish populations

Introduction

The aim of this study is to identify and estimate population structuring in alewife populations using microsatellite loci. The first chapter deals with background information on population genetics and the factors and mechanisms underlying population differentiation in natural populations with emphasis on fish populations and their management implications. Also included is a small section on molecular markers such as microsatellites, and their use in population genetics. The second chapter, which presents my work, is prepared as a manuscript for submission and the final chapter discusses the interpretation of my data in a larger context of implications in management of fish populations.

Genetic variation in natural populations

Natural populations that have restricted gene flow as a consequence of adaptive divergence often exhibit population structuring. Identification of population structuring is intended to help maintain genetic variability in declining populations (Hinten et al, 2003). Maintaining genetic variability in such populations is the primary concern of conservation biology. This is because higher genetic variation is assumed to improve the fitness of individuals and enhance the probability of population survival (Zoller et al, 1999).

Genetic variation is an important factor in the process of evolution in natural populations. Genetic variation arises through mutations and is acted on by the forces of migration, genetic drift and various types of natural selection (Gunderina, 2003). Although in most natural populations all these factors operate simultaneously, genetic drift is said to play a dominant role in determining the level of genetic variation in small and isolated populations. Genetic drift is brought about by changes in allele frequencies and is faster in small populations than in larger ones (Hinten et al,

2003). While genetic drift has a role in small populations, natural selection plays an important role in determining genetic variation in large populations.

Natural selection and adaptive divergence

Populations exist in dynamic environments which are heterogeneous in many dimensions. Under natural selection, populations may confront these fluctuations through phenotypic variations either (i) within a single individual; (ii) among different individuals at one time; or (iii) in future generations. A single organism can display multiple strategies, coping with fluctuations in different facets of the environment (Meyers and Bull, 2002). Nonetheless, the minimum requirement for an evolutionary change to occur under natural selection is the presence of heritable variation in the selected trait (Hoffmann and Merila, 1999).

Natural populations usually have a range of geographic distribution and are exposed to different environments in different locations. Under differential selection, individuals tend to adapt to their local environmental conditions resulting in a pattern of local adaptation (Lenormand, 2002). Also, depending on the local selective pressures, populations exposed to different ecological environments diverge for phenotypic traits which influence their survival and reproduction (Hendry, 2001). As a result of this adaptive divergence, migrants from other environments would become less fit than local residents in a particular environment. This could prevent interbreeding and hence reduce gene flow between populations and may eventually lead to ecologically-dependent reproductive isolation (Hendry, 2001). Hence reproductive isolation may evolve as a consequence of phenotypic divergence due to natural selection (Smith et al, 2001).

Gene flow and selection

Gene flow is said to keep a check on divergence by opposing the effects of natural selection (Lenormand, 2002). This is illustrated by the fact that populations connected by high levels of gene flow are usually less differentiated (McKay and Latta, 2002). This occurs because gene flow tends to homogenize populations by swamping local adaptation and preventing

differentiation caused by selection. When migration rate, or in other words gene flow, is large compared to local selection, the alleles that confer the best average reproductive success across all populations tend to become fixed (Lenormand, 2002) hence compromising local adaptive evolution (Storfer, 1999). In fact, population genetics theory suggests that one migrant per generation is sufficient to prevent population differentiation in the absence of natural selection (Speith, 1974). This theory was originally developed for "island" models where very low migration and mutation rates can disrupt genetic differentiation. However, this theory is merely a "rule of thumb" and is applicable to populations under negligible selection pressures (Speith, 1974).

Conversely, in the presence of natural selection, populations can develop adaptive differences in spite of considerable gene flow since strong selection can remove the genetic load imposed by immigrants, maintaining differences among populations (McKay and Latta, 2002). For instance it was shown that two populations of sockeye salmon which were subjected to differential selection were genetically distinct even though they interbreed to some degree, with an average of about 39% immigrants into the river in which they breed (Hendry, 2001). Reproductive isolation due to adaptive divergence is also said to evolve in a short amount of time. It was shown experimentally that ecologically-dependent partial reproductive isolation evolved in about 13 generations in two populations of sockeye salmon, demonstrating that adaptive divergence can quickly lead to partial reproductive isolation. It was also shown that morphological differences in males and females and differences in the development of embryos between two river populations lead to reproductive isolation. Using single trait equations for selection and calculation of expected distribution of male and female body lengths based on random genetic drift, it was shown that natural selection was the mechanism leading to divergence in these populations (Hendry, 2001).

Reproductive isolation can develop rapidly in allopatric, parapatric or sympatric populations when divergent selection is strong relative to gene flow. Allopatric populations are those which are separated by geographic barriers, parapatric are neighboring populations along ecological gradients and sympatric populations are those which are separated by ecological barriers. In allopatry, adaptation to geographically distant and differing environments contributes

to gradual genetic divergence. Gene flow in these populations is restricted due to geographic barriers allowing genetic differences to accumulate, leading to reproductive isolation (Tregenza, 2002). In parapatry, natural selection can promote phenotypic or morphological divergence across ecological gradients leading to reproductive isolation (Smith et al, 2001). In sympatry, it was shown that multifarious, strong divergent selection coupled with pleiotropy could lead to complete reproductive isolation (Wood and Foote, 1996). Sympatric genetic divergence of benthic and limnetic morphs of sticklebacks is an example of such reproductive isolation. Benthic and limnetic morphs differ in their morphology, habitat use and foraging ecology in the lakes in which they occur. Pre-mating isolation seems to occur due to differences in nesting locations and assortative mating by body size and coloration. Post zygotic isolation seems to occur due to natural and sexual selection against hybrid genotypes. Both pre-zygotic and post-zygotic isolation mechanisms lead to genetic divergence in these morphs (Taylor, 1999).

Fish biology

Fishes are the largest group of vertebrates in terms of number of species (McLean et al, 1999) with wide ranges of geographical distribution. A behavioral character complex among fish is diadromy, which is characterized by migrations between fresh water and the sea. The migrations are regular, physiologically mediated movements between the two biomes, occurring at characteristic phases of life history in each species. They are usually obligatory and necessarily involve two reciprocal migrations since they constitute a species' life cycle (McDowall, 1997).

Life history strategies

Diadromy can be divided into three categories; anadromy, catadromy and amphidromy (table 1). About 227 species were listed as diadromous, out of which 110 were recognized as anadromous (McDowall, 1997). Anadromous species occupy multiple freshwater, estuarine and marine habitats. Anadromous fish spend their early life stages (egg-juvenile) in freshwater and occupy marine environments during their adult life stages returning to freshwater to spawn

(Quattro et al, 2002). When the lakes to which the anadromous fish migrate become land-locked, preventing migrations to and from the sea, landlocked populations abandon diadromy and eventually evolve into distinct daughter species. The family Clupeidae includes such non-migratory species derived from diadromous species (McDowall, 1997).

Table 1: Three categories of diadromous migrations (McDowall, 1997).

Anadromy	Feeding and growth take place in the sea and fully grown adult fish migrate to fresh water to reproduce.
Catadromy	Feeding and growth take place in freshwater and fully grown adult fish migrate to sea to reproduce.
Amphidromy	Migration of larval fish to sea soon after hatching, early feeding and growth at sea and then migration of small juvenile fish from sea to freshwater for prolonged feeding, sexual maturation and reproduction.

The pre-requisite for the evolution of diadromy seems to be the ability to tolerate and osmoregulate across a wide range of salinities (euryhaline). This seems to be a common feature among diadromous fish (McDowall, 1997). Prior to seawater migration, anadromous fish like salmonids undergo a pre-adaptive transformation, which involves morphological, biochemical, physiological, and behavioral alterations (Jorgensen and Arnesen, 2002). Anadromous Arctic charr must adapt to rapid changes between hyper-osmotic and hypo-osmotic regulatory conditions in order to keep blood plasma osmolality and mineral levels within acceptable limits (Gulseth et al, 2001). External factors like photoperiod and temperature seem to be essential cues for the development of hypo-osmoregulatory capacity in these fish. Pre-migratory improvement of the hypo-osmoregulatory capacity is accompanied by an increase in the gill Na^+, K^+- ATPase activity (Jorgensen and Arnesen, 2002). The gill chloride cells exhibit high functional activity and the kidneys have nephrons with two proximal segments, which are capable of excreting large amounts of magnesium. Both the kidneys and the intestine are shown to be involved in the excretion of excess ions and maintenance of serum hypo-osmolarity in all the species studied (Krayushkina et al, 2001). Seawater adaptability within a population was also shown to vary with size of the fish (Gulseth et al, 2001).

Population genetics

Fishes display various degrees of among-population differentiation with intra-specific diversity partitioned with respect to geography. The number and locations of genetically distinct populations within a species vary depending on environmental factors and life-history traits (McLean et al, 1999). Life history strategies including dispersal capacities play a significant role in determining genetic variability and population structure in fish species (McLean et al, 1999). The potential for dispersal of fishes can be divided into three life history categories: fresh water, anadromous and marine. These determine the capacity of gene flow among populations and its reflection in the genetic differentiation of populations (McLean and Taylor, 2001). Riverine and lacustrine fish populations are highly structured by geography with deep population structure observed between adjacent drainages (Quattro et al, 2002). This could be because geographical discontinuities in the distribution of habitat might limit the exchange of fish and their genes among locations (Youngson et al, 2003). Coastal marine species show shallow population structure on a much broader geographic scale (Quattro et al, 2002). Though the lack of physical barriers in the ocean allows a great degree of mixing between fish from different locations, marine fishes do exhibit some genetic structure. Behavioral limits to dispersal are among the various factors responsible for population subdivision in marine species (McLean et al, 1999). Anadromous species, on an average have an intermediate level of genetic structuring, which usually depends on the levels of straying (Quattro et al, 2002).

In general, many fish species comprise morphologically, ecologically and genetically distinct populations that are sympatric during at least some portion of their life cycle. Such sympatric fish species contribute to the biodiversity of temperate fresh water ecosystems (Taylor, 1999). Non-anadromous populations have the potential to develop river-specific local adaptations and become distinct from anadromous populations (Tessier et al, 1997). Wood and Foote (1996) showed that the anadromous and non-anadromous morphs of sockeye salmon spawning in the same rivers were genetically distinct. Differences in morphological, developmental and reproductive traits seem to promote reproductive isolation between these two morphs. Differences in characters such as size, number of gill rakers and age at maturity are associated

with ecological differentiation and are heritable in sockeye salmon (Wood and Foote, 1996). These differences between anadromous and non-anadromous forms are likely to exert strong divergent selection between the two populations. This kind of divergent selection, strong assortative mating (based on size differences), and choosing different microhabitats for spawning contribute to reproductive isolation in sympatry (Taylor, 1999). Such prezygotic isolating mechanisms appear to reduce gene flow among morphs and promote genetic differentiation between them (Wood and Foote, 1996). Such ecological reproductive isolation among lacustrine populations living in sympatry is a common feature of many northern fishes (Tessier et al, 1997).

Alewife (*Alosa pseudoharengus*), blueback herring (*Alosa aestivalis*) and American shad (*Alosa sapidissima*) belong to the family Clupeidae, commonly called the herring family. This family includes mostly anadromous, and a few catadromous or amphidromous fish, although the life history strategy is consistent within genera (McDowall, 1997). These anadromous herrings show a high degree of flexibility and intrageneric niche differentiation in places where they co-occur and are reproductively isolated species (Munroe, 2002). Because of their unique life history, the population genetics of anadromous fish is of particular interest.

Anadromous fish reproduce in freshwaters where juveniles spend a few years before migrating to the marine feeding grounds. In most cases, after a few years of growth in the ocean, the sexually mature fish return to their natal sites to reproduce and complete the life cycle (Tessier et al, 1997). Iteroparous species, which reproduce multiple times in their life span might exhibit two kinds of reproductive fidelity, spawning-site fidelity and natal-site fidelity, while semelparous species, which reproduce only once in their life span show natal-site fidelity. Spawning-site fidelity occurs when individuals return to the same spawning grounds in consecutive reproductive seasons and natal-site fidelity occurs when individuals return to the spawning grounds of their birth (Miller et al, 2001). Anadromous fish usually show high site fidelity for reproduction, which is referred to as philopatry or homing behavior. Philopatry is also observed in other animals like wood-rats (Matocq and Lacey, 2004) and seabirds (Steiner and Gaston, 2005).

High degrees of philopatry or homing behavior has been shown in many fish species including but not limited to the eulachon (McLean et al, 1999; McLean and Taylor, 2001), Atlantic salmon (King et al, 2001), European hake (Lundy et al, 1999), striped bass (Brown et al, 2003), brown trout (Bouza et al, 1999), redband trout (Nielsen et al, 1999) and American shad (Waters et al, 2000). American shad is a good example of a species that shows natal site fidelity, which was estimated to be as high as 97% in the Annapolis River (Waters et al, 2000). Such fidelity can lead to genetic differences and structuring among spawning populations resulting from reproductive isolation (Miller et al, 2001).

The life history of anadromous fish favors the formation of high numbers of local populations exhibiting a large degree of differentiation across contrasting environments (Tessier et al, 1997). These populations are susceptible to the effects of natural selection and, under many circumstances, show local genetic adaptation (Youngson et al, 2003). In addition, reproductive fidelity can lead to isolation among populations and consequently to genetic differentiation. Spawning-site fidelity and natal-site fidelity ensuring return to their site of birth for spawning in subsequent reproductive seasons would restrict inter-breeding or gene flow among populations and allows genetic differences to accumulate (Miller et al, 2001). Hence, spawning fidelity might lead to progressive genetic divergence among different river populations with time (Brown et al, 1999). In other words, homing to their natal rivers, repeated over generations, limits gene flow and allows the accumulation of genetic differences among anadromous fish populations (McLean et al, 1999). Genetic structuring in such populations is usually measured using molecular markers and related statistics.

Molecular markers

Analysis with molecular markers has proven to be a strong method of identifying genetic differentiation among populations and population structuring (King et al, 2001). Some of the molecular markers generally used in investigating genetic variation are allozymes, mitochondrial DNA, microsatellites and minisatellites.

Some of the advances in molecular biology which increased the efficiency of these markers include development of the polymerase chain reaction (PCR) which amplifies specified stretches of DNA, and the advent of routine DNA sequencing (Sunnucks, 2000). PCR-based markers offer many advantages such as the use of minute tissue samples, large scale automation and no need for membrane hybridization (Mitchell-Olds, 1995). Table 2 compares advantages and disadvantages of some common molecular tools (O'Connell and Wright, 1997). Because of the advantages like low screening costs and the high potential for estimating population structuring compared to other molecular markers, microsatellites were chosen for the present study.

Table 2: Comparative advantages and disadvantages of some commonly applied molecular tools (table from O'Connell and Wright, 1997).

Technique	Technical requirements		Cost		Storage requirements	Potential for		
	Development	Screening	Development	Screening		Genome mapping	Parentage assessment	Population genetics
Allozymes	Low	Low	Low	Moderate	High	Low	Low	Moderate/high
MtDNA[1]	Low	Low	Nil	Moderate	Low	Nil	Low/moderate	Moderate/high
RAPDs[2]	Low	Low	Nil	Low	High	High	Moderate	Low
ScnDNA[3]	High	High	High	Moderate/high	Moderate	High	Low	Moderate
ESTs[4]	High	High	High	Moderate	Moderate	High	Low	Low
MLF[5]	Moderate	Moderate	Low	High	Moderate	Nil	High	Low
SLPs[6]	Moderate/high	High	High	Moderate	Moderate	High	High	High
SSRs[7]	Moderate/high	High	High	High	Low	High	High	High
SSRs[8]	High	Low	High/moderate	Low	Low	High	High	High

[1]PCR-amplified mitochondrial DNA.
[2]Randomly amplified polymorphic DNA.
[3]Single-copy nuclear DNA restriction fragment length polymorphisms.
[4]Expressed sequence tags.
[5]Multilocus fingerprinting.
[6]Single-locus minisatellite probes.
[7]Dinucleotide simple sequence repeats (microsatellites).
[8]Tri-, tetranucleotide simple sequence repeats (microsatellites).

Microsatellite markers

Typically, a diploid organism has one or two different states (alleles) per character (locus). Variable genetic markers reflect differences in DNA sequences, measuring the allelic variation among populations. Microsatellites are single-locus markers which are flexible and informative because they can be analyzed as genotypic arrays, as allele frequencies (Sunnucks, 2000). Microsatellites are simple sequence repeats (SSRs) consisting of short tandem arrays of

repeated units. These SSRs are highly abundant and dispersed throughout the euchromatic part of the genome and thought to occur every 10kbp of the genome in fish species (O'Connell and Wright, 1997). Depending on the length of the repeat unit, microsatellites are classified as mono-, di-, tri-, penta- and hexa-nucleotide repeats. Microsatellite stretches may be disrupted by base substitutions (imperfect) or insertions (interrupted) or might consist of more than a single repeat type (compound microsatellite). Examples for these terms are given in figure 1 (Schlotterer and Harr, 2001).

Figure 1: Microsatellite terminology (from Schlotterer and Harr, 2001).

Mononucleotide: $(A)_{13}$	AAAAAAAAAAAAA
Dinucleotide: $(GT)_8$	GTGTGTGTGTGTGTGT
Trinucleotide: $(GAT)_7$	GATGATGATGATGATGATGAT
Tetranucleotide: $(CTAG)_6$	CTAGCTAGCTAGCTAGCTAGCTAG
Pentanucleotide: $(CATTG)_5$	CATTGCATTGCATTGCATTGCATTG
Hexanucleotide: $(GGATCC)_4$	GGATCCGGATCCGGATCCGGATCC
Imperfect microsatellite:	GTGTGTGTGTATGTGTGTGTG
Interrupted microsatellite:	GTGTGTGTCCCGTGTGTGTGT
Compound microsatellite:	GTGTGTGTGCTCTCTCTCTCTC

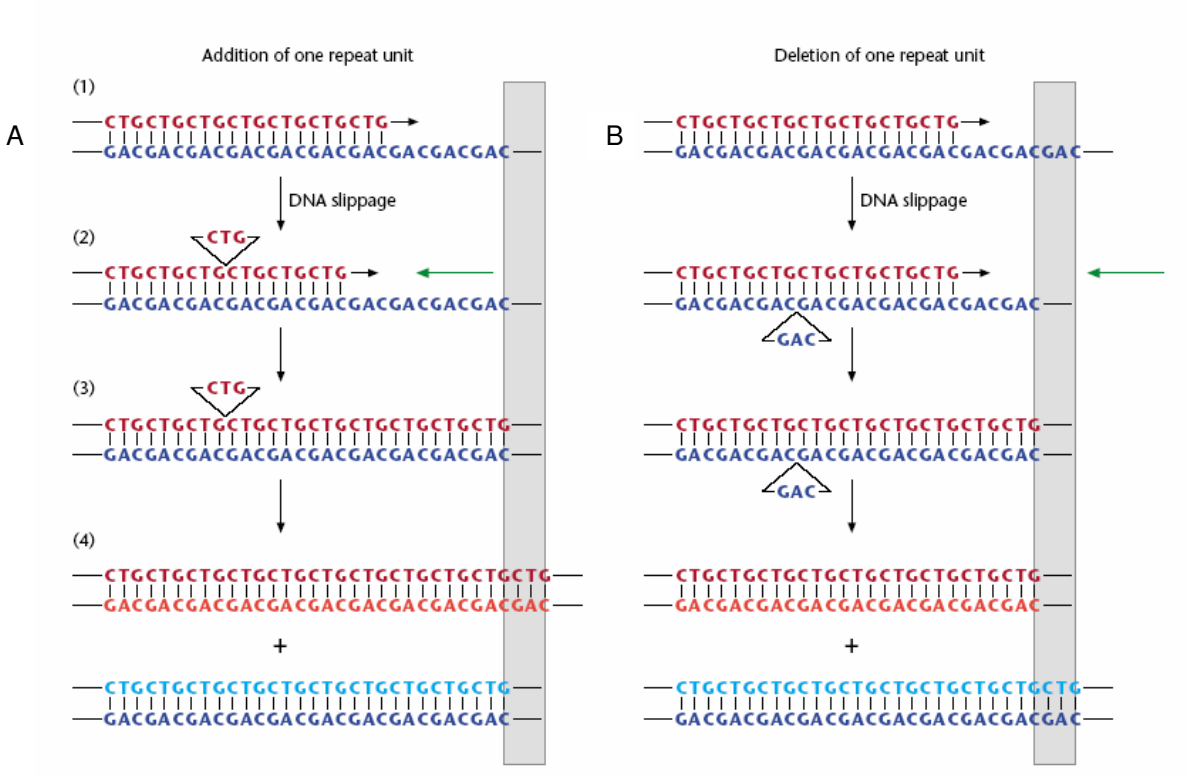
Microsatellites may be found within expressed regions of the genome. Since limitation on allele size among microsatellites has been detected, these loci are thought to be under selective pressure (Jarne and Lagoda, 1996). Microsatellites have become the preferred marker in many studies because of their high levels of variability and co-dominant inheritance. They also generally provide precise and accurate estimates of population parameters such as effective population size and intra- and inter-locus disequilibrium (Luikart and England, 1999). Microsatellite assays are also PCR-based which means minute quantities of tissue can be used for the assays (O'Connell and Wright, 1997).

The mutation process in microsatellites occurs through slippage replication, in which loci accumulate or lose repeat units due to a proportion of primary slippage events that escaped the

mismatch repair system (Schlotterer and Harr, 2001). Microsatellite loci are usually non-coding and hence tend to escape the mismatch repair system, which is more efficient in the coding parts of the genome. During DNA replication, the newly synthesized strand tends to get displaced and realigns out of register. Since microsatellite DNA has repeat sequences, both the strands can still pair, leaving a small loop. When the DNA replication continues without repairing the loop structure, this kind of slippage results in the gain or loss of repeat units depending on which strand forms the loop structure. Figure 2 explains the microsatellite mutation model (Schlotterer and Harr, 2001). Mutation rates in microsatellites range from 10^{-5} to 10^{-2} per generation (Jarne and Lagoda, 1996).

Figure 2: Model of DNA slippage

A) adding or B) removing one repeat unit. (1) First round of DNA replication. (2) DNA slippage, causing one repeat unit to loop out. (3) DNA synthesis continues without repair of the loop. (4) Second round of DNA replication leads to addition or deletion of one repeat unit in one of the DNA strands (figure from Schlotterer and Harr, 2001).



Understanding the mutational model underlying microsatellite evolution is of great importance for developing statistics for analyzing microsatellite data. Two models developed for mutations of alleles include the infinite alleles model (IAM) and the stepwise mutation model (SMM). In the IAM, each mutation gives rise to a novel allele at a given rate. Because every mutation results in a unique allele, identical alleles are assumed to share a common ancestry and are therefore considered to be identical by descent (IBD). Under the SMM, each mutation gives rise to a novel allele either by adding or deleting repeated units, with equal probability in both directions. Hence identical alleles need not be identical by descent (Balloux and Moulin, 2001). It is generally thought that microsatellites predominately mutate by one or a few repeats only and therefore the SMM is often used to describe their mutational process (Schlotterer and Harr, 2001).

Homoplasy, the co-occurrence of alleles that are identical by state though not identical by descent (Jarne and Lagoda, 1996), is expected to occur under SMM (Estoup et al, 2002) but not under IAM. Microsatellite alleles or electromorphs identical in state would be identical in size, but not necessarily identical by descent due to the possibility of convergent mutations. Hence, homoplasy occurring at microsatellites is referred to as size homoplasy (Estoup et al, 2002). Size homoplasy reduces the observed number of alleles per population, the proportion of heterozygous individuals and genetic diversity (Estoup et al, 2002). Although it is theoretically true, Rousset (1996) has shown that there is no effect of homoplasy on the parameters F_{IS} and R_{IS} and no simple effect on parameters F_{ST} and R_{ST} (Estoup et al, 2002). Furthermore, Adams et al, 2004 have shown that while 7 of 12 loci demonstrated homoplasy for a tropical tree *C.alta*, there was no homoplasy detected for any of the 11 loci for an anadromous fish, *M.saxatilis.* They also found that the effects of homoplasy were of the same order of magnitude as the sources of data error in studies utilizing microsatellite loci. Hence homoplasy might not be of much concern in population genetic studies.

Fixation indices

Fixation indices F_{ST} and R_{ST} are the most common statistics reported for estimation of population structure. Other F-statistics used for populations divided into sub-populations include F_{IS}, F_{ST} and F_{IT} where I, S and T stand for individuals, sub-populations and total populations respectively (Wright, 1965). F_{ST} is defined as the correlation between two alleles chosen at random within subpopulations relative to alleles sampled at random from the total population. F_{ST} reaches a value of one when the two subpopulations are totally homozygous and fixed for alternative alleles and a value of zero when allele frequencies in the two subpopulations are identical (Balloux and Moulin, 2001). Table 3 shows the suggested genetic differentiation for different ranges of F_{ST}.

Table 3 Interpretation of F_{ST} values (Balloux and Moulin, 2001).

Value of F_{ST}	Extent of genetic differentiation
0-0.05	Little
0.05-0.15	Moderate
0.15-0.25,	Great
above 0.25	very great

F statistics can be used for almost all species, but the accuracy of interpretation depends on factors like sample size, number of populations or number of alleles etc. Though there is some disagreement about the interpretation of the F values, the interpretation given in table 3 will be used in this study.

As mentioned, population genetic structure measured with different molecular markers is frequently quantified using F_{ST} (Wright, 1965) or related statistics, which measure the proportion of total allelic variation that occurs between populations (Balloux and Moulin, 2001). Nielsen et al (1999) estimated genetic differentiation of redband trout populations from paired comparisons of mainstream and tributary populations of McCloud River, California using R_{ST}, an analogue of F_{ST} that may be more appropriate for analysis of microsatellite markers. The values ranged from

0.002 to 0.607 for these populations (Nielsen et al, 1999). The average F_{ST} values reported for some of the fish species using different kinds of molecular markers is given in table 4 (O'Connell and Wright, 1997).

Table 4: Average estimates of differentiation using F_{ST} or analogous differentiation coefficients, reported for different marker systems in a number of fish species (table from O'Connell and Wright, 1997).

Species	Allozymes	NRFLPs	mtDNA	Minisatellite DNA	Microsatellite DNA
Atlantic cod, *Gadus morhua*	0.014[1]	0.069[1]	NS[2*]	0.030[3]	0.035[4]
Atlantic salmon, *Salmo salar*	0.064[5,6,7]	No data	0.106[5]	0.036[8]	0.028[9,10]
Rainbow trout, *Oncorhynchus mykiss*	0.080[11]	No data	Not reported	0.049[12]	0.016[13]
Brown trout, *Salmo trutta*	0.190[14]	No data	0.309[14]	0.174[14]	No data
Pacific herring, *Clupea harengus*	0.014[15]	No data	No data	No data	0.023[16]
Brook charr, *Salvelinus fontinalis*	0.232[17]	No data	0.128[17]	No data	0.515[18]

[1]Pogson *et al*. (1995), based on six North Atlantic populations.
[2]Arnason *et al*. (1992), based on 20 populations from Iceland and North-west Atlantic populations.
[3]Galvin *et al*. (1995a), based on five North-east Atlantic populations.
[4]Bentzen *et al*. (1996), based on six North-west Atlantic populations.
[5]O'Connell *et al*. (1995), based on seven tributary populations from three drainages.
[6]Stahl (1987), based on 31 populations from throughout the North Atlantic.
[7]Jordan *et al*. (1992), based on 22 tributary populations from 16 drainages.
[8]Galvin *et al*. (1995a), based on nine tributary populations.
[9]O'Connell *et al*. (1996b), based on five tributary populations.
[10]McConnell *et al* (1997a), data based on 15 populations.
[11]Hershberger (1992), based on 38 populations from western North America.
[12]Taylor (1995), based on eight north-eastern Pacific populations.
[13]O'Connell *et al*. (1996a), based on five Lake Ontario populations.
[14]Ferguson *et al*. (1995), based on 29 populations for allozymes, 25 populations for minisatellites and 14 populations for minisatellites.
[15]Grant and Utter (1984), based on 21 North Pacific populations.
[16]O'Connell *et al*. (1996b), based on seven North-east Pacific populations.
[17]Jones *et al*. (1996), based on 34 Eastern Canadian populations.
[18]Angers *et al*. (1995), based on five populations from two drainages.
*NS, not significant.

F_{ST} was developed assuming the infinite alleles model, but since microsatellites seem to follow a stepwise mutation model, an analogue R_{ST} was developed. R_{ST} is calculated from the variances in allele sizes, whereas F_{ST} is derived from variances in allele frequencies (Balloux and

Moulin, 2001). The relative performance of these two statistics depends on various factors which cannot be quantified and hence a careful interpretation of both statistics might give valuable information about the genetic structure of populations (Balloux and Moulin, 2001).

Population genetics and fisheries policy

The economic value of recreational fisheries derives from factors such as overall abundance of individuals, overall genetic diversity and size of geographical range. These factors affect their availability across local fisheries and also throughout the seasons (Youngson et al, 2003). Hence, it is beneficial to pay attention to genetic as well as demographic variables for biological restoration of viable populations (Brown et al, 2000).

Recreational fishes are often commercially exploited and a great number of fish, including freshwater and anadromous fish, are seriously threatened as a consequence of human activities (Madeira et al, 2005). A number of factors may have contributed to the decline of anadromous fish populations. The decline in the population size of sturgeon, allis shad (*Alosa alosa*) and twaite shad (*Alosa fallax*) could be attributed to over-fishing, destruction of spawning habitats and pollution (Groot, 2002). Canalization and construction of dams and weirs preventing migration of fish to their spawning habitats is another potential factor and has caused declines in sturgeon and smelt (*Osmerus eperlanus*) populations (Groot, 2002). Alewife populations have been excluded from many inland waters of Maine, USA and elsewhere due to dams and pollution (Moring and Mink, 2002). Pollutants may also interfere with the development of olfactory imprinting (Groot, 2002), which can impair homing and spawning behavior and thus might effect their reproductive success (Scholz et al, 2000). Substantial differences in relative abundance of juvenile alewife and blue-back herring have been reported in coastal streams of southern New England (Kosa and Mather, 2001). Culverts, bridges and low-head dams are common obstacles in many small coastal New England streams and can modify emigration patterns of juvenile alewife across systems (Kosa and Mather, 2001). Closure of sea inlets has caused the decline of coregonids and sea trout populations in the Rhine, Netherlands. Increases in silt levels caused by sand and

grave extraction, combined with pollution have caused salmon populations to decline. The increase in silt levels in the river and subsequent sedimentation on the spawning grounds has made these sites unsuitable for reproductive purposes (Groot, 2002).

Regional fishery management councils prepare management plans for US coastal fish stocks and they are under the supervision of federal agencies like the National Marine Fisheries Service. They are responsible for biodiversity, conservation and survival of fish species (Corkett, 2005). Restocking, one of the strategies used by such agencies is defined as repeated injection of fish into an ecosystem in which the species is already native to that water body or exotic to it but previously introduced (Aprahamian et al, 2003). Restocking when used, helps restore, enhance and conserve populations within the scope of natural production. For example, rehabilitation efforts of Atlantic salmon in the Meuse basin, Belgium gave promising results with respect to salmon survival, growth rates, smoltification and migration patterns. Young salmon in this area also showed good adaptation in tributaries where they were implanted, with parr (early juvenile stage) densities up to 20-30 individuals per $100m^2$ (Prignon et al, 1999).

While restocking can benefit the commercial and recreational fisheries on a short term, there are also risks associated with it. It might impact other organisms that have conservation value or it might give rise to competition and/or predation (Aprahamian et al, 2003). For example, reintroduction of alewife in Maine that had white perch populations affected both white perch and alewife juveniles. This was caused due to predation by adult alewife on larval stages of white perch, predation by adult white perch on the juvenile alewife population and also by cannibalism from adult alewives (Moring and Mink, 2002). It has also been shown that restocking programs can result in deleterious effects on the natural fish populations, by increasing disease and competition for food and habitat which might result in their extinction (Madeira et al, 2005). The main concern, however, is the potential genetic impact on natural populations (Aprahamian, 2003). For example, overall loss of genetic variation with about 27% decrease in heterozygosity was observed in Atlantic salmon in the Tasmanian aquaculture industry compared to the parent Canadian populations (Reilly et al, 1999). Hence, restoration efforts should take into account inter- and intra-river genetic diversity and should avoid significantly perturbing the recipient

population caused by shifting gene frequencies, influencing demographic and physiological parameters or introducing disease (King et al, 2001). Molecular genetic markers guide in short-term actions such as transfer of individuals between existing populations as well as restoration efforts where populations have been extirpated (McKay and Latta, 2002).

Maintaining genetic diversity among different sites safeguards against extinction caused by environmental perturbations or diseases (Uthicke and Purcell, 2004). Introduction of fish of the same species from a different geographical origin might bring about genetic changes in native populations (Madeira et al, 2005) in such a way that alleles introduced by these fish could replenish local genetic variation (Lenormand 2002). This genetic variation affects fitness in wild populations adapting to different environments (Mitchell-Olds, 1995).

However, locally adaptive genetic diversity within units is likely to be of greater importance when choosing populations that are more suitable as translocation or restoration sources. Adaptive genetic differences among populations can lead to outbreeding depression if divergent populations are mixed (McKay and Latta, 2002). Outbreeding depression occurs due to swamping of locally adapted genes by genes from exotic populations or selection against hybrids, which cannot perform well in either environment. Genetic introgression might also cause homogenization of wild populations from different geographical areas, which might lead to the loss of local adaptation and the decline of natural self-sustaining populations. Anthropogenic hybridization might reduce the genetic variability between native populations and populations in other geographic areas from where fish are being introduced and hence constrain future adaptations (Madeira et al, 2005; Marzano et al, 2003). Another danger of change in genetic diversity between populations would be if it affected the fitness of individuals in a way that it resulted in changes in the transferred stock, making it unable to adapt into the new habitats (Nielsen et al, 1999).

For example, brown trout populations have been depleted in many natural environments. Artificial selection of domestic stocks coupled with over-fishing has been responsible for the loss of genetic diversity in these populations. Massive restocking with domesticated or hatchery strains obtained from few spawners has shown to result in “founder effects” due to low genetic

diversity of spawners (Marzano et al, 2003). Another example is of Atlantic salmon. A supportive-breeding program was initiated for restocking Atlantic salmon in which a fraction of wild fish were brought into a hatchery for artificial reproduction and the offspring were released into the river of origin to interbreed with the wild fish. Such a practice is predicted to have the potential to result in increased inbreeding, reduction in heterozygosity, and loss of rare alleles all leading to detrimental effects on wild populations. This mainly depends on the number of captive spawners and their relative contribution to overall progeny (Tessier et al, 1997).

The success of stocking efforts should therefore be judged on the reproductive potential of the stocked fish and their subsequent contribution and/or impact on the native population. A crucial component of recovery plans is conservation of the composition and pattern of genetic diversity within populations of endangered species (Quattro et al, 2002).

Conclusion

Alewife, fish of economic value have been declining and restocking efforts are being put in place to protect these populations. Since alewife are anadromous fish and expected to have population structuring and since maintaining genetic diversity is an important and critical aspect in conserving the declining populations, identification of genetic structuring in alewife might assist in management programs. The following study involves estimation of genetic structuring in alewife populations using microsatellite markers.

CHAPTER 2

Genetic differentiation in alewife populations using microsatellite loci

Introduction

Populations of many species show some level of genetic structuring which may be due to factors like environmental or ecological barriers, historical barriers or life histories (McLean et al, 1999; Lundy et al, 1999). Anadromous fish spend most of their life in the sea and enter fresh water to spawn. Spawning-site fidelity should result in reduced gene flow among populations and hence might lead to progressive genetic divergence among different river populations with time (Brown et al, 1999). Genetically structured breeding populations may have different effective population sizes, distinct reproductive rates and variable susceptibility to harvesting or breeding habitat degradation (Brown et al, 2003).

Alewife, *Alosa pseudoharengus,* are anadromous fish found along the Atlantic coast from Newfoundland to North Carolina (Daniels, 2001). Adult alewife migrate in early spring from marine ecosystems to fresh waters to spawn (Limburg, 1998). They spend most of their life as mixed populations in the ocean but are highly philopatric at the time of spawning (Munroe, 2002). Olfactory clues appear to play an important role in the high rate of fidelity with which they return to their natal streams to spawn (Munroe, 2002). Landlocked populations of alewife also occur in the Great Lakes (Moring and Mink, 2002).

It has been shown that energy invested in reproduction by anadromous fish is negatively correlated with the distance of upstream migration (Wood and Foote, 1996). Alewife populations in southern New England show physiological differences with respect to the location at which they spawn (E.Schultz, University of Connecticut, personal communication, Feb 2004). Female and male alewife running up Bride Brook (which drains directly into Long Island Sound) have larger gonads than alewife running into Roaring Brook (which drains into the Connecticut River, figure 3). There are also age differences among fish running in the two streams; the Bride Brook population's modal age is 3 yrs and the Roaring Brook population's modal age is 4 yrs.

Alewife have both recreational and commercial value (Munroe, 2002). They are used as bait for commercial fisheries including crab and lobster. Their eggs are canned and eaten as a delicacy and they are used as pet food. But perhaps more importantly, they are a forage base for sport fish like striped bass and blue fish, which support a robust recreational industry. Alewife commercial landings in New England have declined significantly over the last forty years from 17,000 metric tons in 1960 to less than 1,000 metric tons in 2002. This has led the Atlantic States Marine Fisheries Commission to declare alewife a species of special concern and to implement conservation measures for their populations. These measures include restocking operations aimed at restoring the alewife populations (Munroe, 2002).

While restocking can benefit the commercial and recreational fisheries on a short term, there are also risks associated with it. The main concern is the potential genetic impact on the populations (Aprahamian, 2003). In anadromous species, differentiation of populations in different rivers may be due to a number of factors which include reproductive isolation due to homing behavior and adaptive responses to different local conditions (Brown et al, 2000). Understanding gene flow, local selection and their effects on populations is useful in many fields of research including conservation biology and population ecology (Wilson et al, 2004). Population genetics studies also aid in identification of different breeding groups and recognizing and protecting genetically distinct and unique populations (Brown et al, 2003). Identification of population structure can assist in designing appropriate restocking programs and therefore help in designing the conservation and management strategies (Brown et al, 2003).

Examination of genetic differentiation using variable microsatellite loci is well established in fisheries, and the high number of alleles makes them sensitive for detecting inbreeding in aquaculture populations (Reilly et al, 1999). Microsatellite loci are genetic markers routinely used in studies of population subdivision, parentage analysis, and shallow phylogenetic relationships and possibly gene flow (Adams et al, 2004). Microsatellites often have a high mutation rate which results in a large number of alleles (Lundy et al, 1999), and therefore renders them highly informative and effective markers for studies of intraspecific population structure (McLean et al, 2001). Microsatellites, which are co-dominant genetic markers, are used to investigate the

genetic structuring of populations in order to address questions of evolutionary and conservation biology.

Here, we use variable microsatellite loci developed for shad (Waters et al, 2000) to examine genetic differentiation of alewife populations in the rivers of southern New England. In many cases, microsatellite loci can be amplified using primers developed for related species (Jarne and Lagoda, 1996) and cross-species amplification was shown in salmonid species (Paterson et al, 2004) and in many other species (D'Amato et al, 1999; Burridge and Smolenski, 2000; Beheregaray and Sunnucks, 2000; Waters et al, 1999).

This project consists of two parts. The first part was to see if the primers developed for shad amplify alewife DNA and the second part to see if there are genetic differences in alewife populations collected from different sites. We show here that microsatellite loci developed by Waters et al. (2000) for shad, *Alosa sapidissima* can be used in alewife, *Alosa pseudoharengus* populations and that preliminary analyses based on small numbers of microsatellite loci suggest that the populations may be genetically different.

Materials and Methods

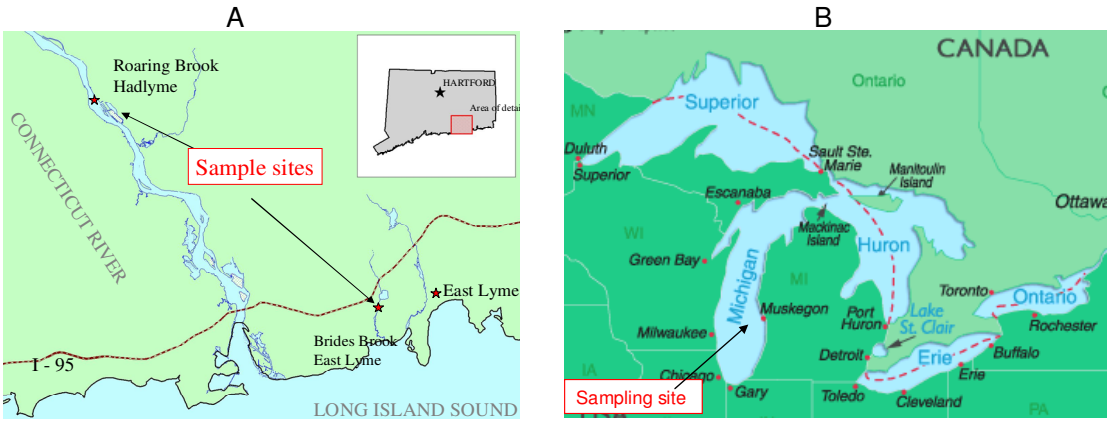
For the first part of this project, DNA was extracted from both alewife and shad and was amplified by PCR using the primers developed for shad. The PCR products were run on agarose gel to look for amplicons which were sequenced to confirm the correct identity. The second part of the project included extraction and amplification of alewife DNA. The amplicons were genotyped and analyzed statistically to identify genetic differences among the different sites sampled. The following section describes these processes in detail.

Sampling

Fish samples were collected from two locations in Connecticut separated by 36.8km and from Lake Michigan. Lake Michigan populations are landlocked descendants of alewife which entered the Great Lakes through the Welland Canal in Ontario, Canada (Daniels, 2001), and made their way to Lake Michigan by 1949 (Madenjian et al, 2004). A total of 56 fish each from Bride Brook and Roaring Brook in Connecticut were collected by weir trapping in March - June of

2003 by Justin Davis (University of Connecticut) and personnel from the Connecticut Department of Environmental Protection Diadromous Fisheries Program. A total of 25 fish were collected from Lake Michigan off Muskegon, MI (figure 3) by trawling during 9/3-4/2003, by Steve Pothoven (University of Michigan) and the crew of the R/V Laurentian. The fish samples were forwarded to our lab on dry ice and 6-10 biopsy punches (5mm each) of the white muscle tissue were collected from each fish and stored at -80°C until further use.

Figure 3 Sampling sites. A) Connecticut and B) Michigan



DNA Extraction

About 0.15g of the white muscle tissue was homogenized in 2ml warm extraction buffer [50mM Tris, pH 8, 0.1M NaCl, 0.1M EDTA ,0.5%SDS] containing Proteinase-K (50ug/ml) for approximately 20 seconds or until the tissue is completely dissociated, using a POLYTRON Tissue homogenizer. DNA was extracted using the phenol-chloroform extraction method (Sambrook et al, 1989). Samples were incubated for 60 min at 55°C in a water bath then treated with 2 ml buffered phenol and rocked for 30 min. Phases were separated by centrifugation and the aqueous phase was removed. Two ml of phenol/chloroform (in a 1:1 ratio) was added to the aqueous phase and rocked for 30 minutes. Phases were separated by centrifugation and the aqueous phase was removed. Two ml of ice cold ethanol (100%) was added and incubated overnight at -20°C to precipitate DNA. The samples were centrifuged at 10,000 g, ethanol was removed and the pellets were allowed to dry completely. The pellets were then re-suspended in 100ul of TE buffer [10mM Tris-Cl, pH 8, 1mM EDTA] and stored at -20°C until further use.

PCR

Six primer pairs developed for shad for the loci Asa2, Asa4, Asa6, Asa8, Asa9, (Waters et al, 2000) and Asa16 (GenBank Accession no. AF049462) were used to amplify the loci in alewife and shad. The 6 microsatellite loci were amplified by polymerase chain reaction (PCR), using the Failsafe system (Epicenter Technologies). Amplifications were carried out in 50µl volumes, in an MJ Research DNA Engine thermal cycler, with initial denaturation at 95^{o}C for 3min, 40 cycles of denaturation at 95^{o}C for one min, annealing at 47-50^{o}C (as specified in Waters et al, 2000), extension at 75^{o}C for 30sec and final extension at 72^{o}C for 10min.

Gel electrophoresis

Loci were examined for successful amplification by running the PCR products on agarose or Metaphor (Cambrex) gels. Loading buffer was added to 10µl of the samples and run on 2% agarose gels using TAE buffer. The microsatellite bands were identified using a 100bp ladder. A 5% Metaphor (Cambrex) agarose was used for the Asa2 locus, since it was difficult to visualize on regular agarose gels because of the smaller size range of these amplicons. A 25 base-pair ladder (Biolone Hyper-ladder V) was used to identify the bands.

Fluorescent Labeling PCR

The samples from alewife that showed bands were again amplified using fluorescently labeled primers and the samples were sent for genotyping to Dr. Tara Paton from the Hospital for Sick Children in Toronto, Canada. Allele sizes were analyzed using the Gene-Mapper software version 3.5 (Applied Biosystems). The software assigns allele sizes based on comparison to a size standard. Allele sizes were then scored by human observers. The smallest allele was assigned the number 1 and other alleles were assigned subsequent numbers based on the kind of repeat. For trinucleotide repeats alleles scored as 3 bases larger than the smallest allele were designated as allele 2 and and alleles scored as 6 bases larger than the smallest allele were designated as allele 3 and so on. Tetranucleotide repeats were also scored in a similar fashion but taking a difference of 4 bases as a new allele. Examples are illustrated in table 5. The original

data and the collapsed data are provided in appendices A and B. Some samples for each locus were also sent for sequence analysis to confirm the identity of the amplicons.

Table 5: Scoring alleles

Alleles for Asa8 (tetranucleotide repeat)	Size in bases
1	116
2	120
3	124
4	128

Data Analysis

Statistical tests were performed using the web version of GENEPOP (Raymond and Rousset, 1995). Tests for Hardy-Weinberg equilibrium and genotypic linkage disequilibrium and probability tests were performed with default values for Markov chain parameters. Hardy-Weinberg tests were performed as exact tests of heterozygote deficit or excess. P values of less than 0.01 were taken to be significant. Genetic diversity was calculated as the number of alleles per locus and observed and expected heterozygosities. The extent of population differentiation was assessed by calculating the fixation indices, F_{ST} (Weir and Cockerham, 1984) and R_{ST} (Slatkin, 1995). R_{ST} is an F_{ST} analogue assuming stepwise mutation model which is thought to be more accurate for microsatellites. The following ranges were used for interpreting F_{ST} values. A value lying between 0-0.05 indicates little genetic differentiation, 0.05-0.15, moderate differentiation, 0.15-0.25, great differentiation and above 0.25 indicates very great genetic differentiation (Balloux and Moulin, 2002). R_{ST} values were also interpreted in a similar manner.

Results

The results indicate that the primers developed for shad do amplify alewife DNA. Except Asa 6, all other loci in alewife were amplified with shad primers. The statistical analyses of the

genotyped data indicate that the alewife populations in the sampled locations may be genetically differentiated.

PCR and sequencing data

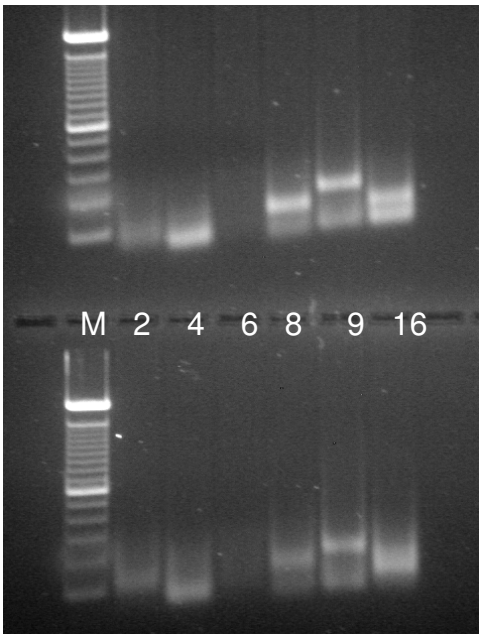
Figure 4 shows the PCR products from amplification of both shad and alewife genomic DNA as template and the six primer pairs developed by Waters et al (2000) for use in shad. As shown, PCR products were obtained for both shad and alewife using five of the six primer pairs (Asa 2,4,8,9 16). The size of the products obtained using alewife DNA were equivalent to those in shad, suggesting that the loci were conserved between these species. However, using primer pair Asa 6, no product was visible with alewife DNA as template.

Figure 4: Agarose gel electrophoresis.

a) loci amplified from alewife DNA b) loci amplified from shad DNA. M=marker, 2=Asa2, 4=Asa4, 6=Asa6, 8=Asa8, 9=Asa9, 16=Asa16.

a)

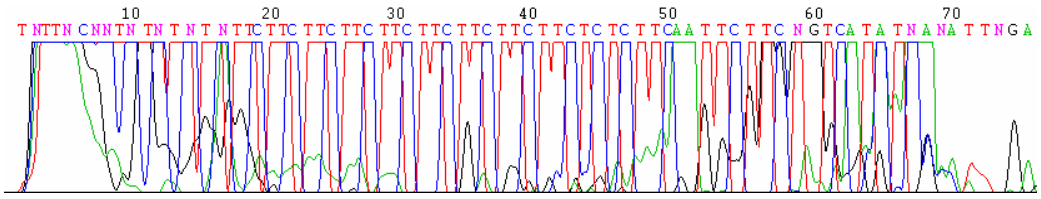
b)



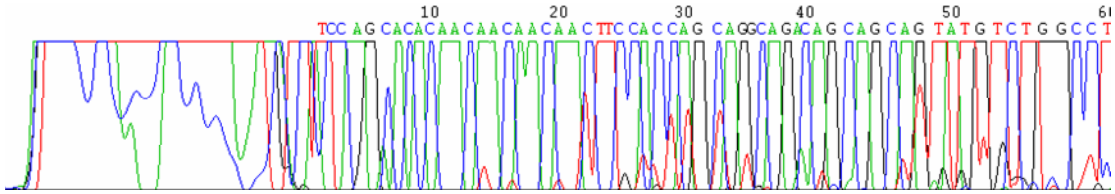
The sequence data from the five loci showed the respective repeat sequences confirming that the targeted microsatellite loci were indeed being amplified and were identical in sequence to those identified in shad. Figure 5 shows the sequence data for each of the five loci.

Figure 5: Sequence data for the five microsatellite loci.

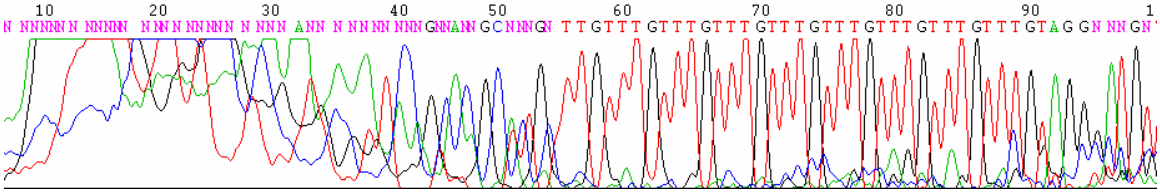
Asa2: $(TTC)_{13}$



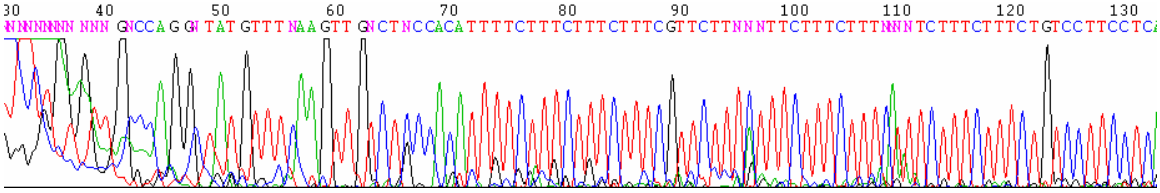
Asa4: $(ACC)_3(AAC)_{12}(AGC)_6$



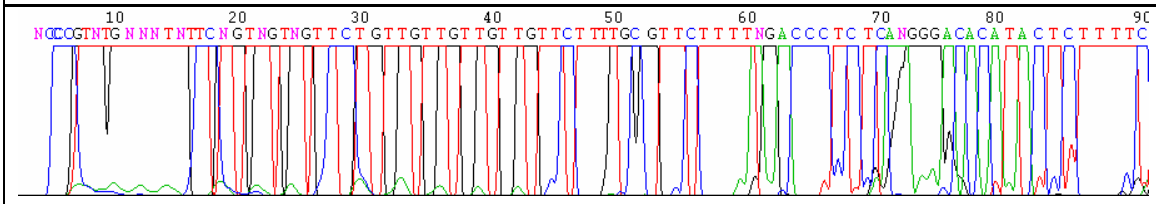
Asa8: $(TTTG)_8$



Asa9: $(TTTC)_7$



Asa16: $(GTT)_3(CCT)(GTT)_{12}$



Allele frequency distribution

All five microsatellite loci showed multiple alleles in alewife, ranging from 6-13 alleles for all populations. One locus, Asa4 showed very low diversity (table 6) and Asa2 did not give any products for Lake Michigan samples. The details of the microsatellite loci with respect to the allele size and number are given in Table 6.

Table 6: Sample size (N), number of alleles, allele size range, and frequency of the most common allele for the five microsatellite loci.
LM=Lake Michigan, BB=Bride Brook, RB=Roaring Brook.

Locus	N			Number of alleles	Allele size range (bp)	Frequency of most common allele (%)
	LM	BB	RB			
Asa2	0	22	23	13	79-148	38.8
Asa4	15	33	43	7	103-129	96.7
Asa8	24	35	36	7	116-139	78.4
Asa9	19	44	50	9	149-202	35.8
Asa16	21	45	52	6	98-127	83.4

Hardy-Weinberg equilibrium and linkage disequilibrium

The observed and expected heterozygosity and the P value for H-W exact test for heterozygote deficit for all populations across all loci are given in table 7. Hardy-Weinberg exact tests were performed using the web-based Genepop software. The observed heterozygosity was in general less than the expected heterozygosity. All populations showed significant heterozygote deficiency at Asa2 and Asa16 loci ($P<0.01$). None of the samples showed significant heterozygote deficiency at Asa4. At Asa8 only Bride Brook population showed significant heterozygote deficiency (P value=0.0068) whereas the other two populations did not. At Asa9, both Roaring Brook and Bride Brook showed significant heterozygote deficiency. No significant heterozygote excess was observed for any population (analysis not shown). A genotypic linkage

disequilibrium test was performed for all populations across all pairs of loci and none displayed significant linkage disequilibrium (analysis not shown).

Table 7: Observed and expected heterozygosity and H-W test (heterozygote deficit). He=expected heterozygosity, Ho=observed heterozygosity, *=No PCR product

Population	Locus	He	Ho	H-W (P-value)
Bride Brook	Asa2	0.650	0.045	0.0000
	Asa4	0.090	0.090	1.0000
	Asa8	0.370	0.310	0.0068
	Asa9	0.800	0.630	0.0005
	Asa16	0.250	0.020	0.0000
	Mean	0.430	0.220	
Roaring Brook	Asa2	0.820	0.400	0.0000
	Asa4	0.020	0.020	-
	Asa8	0.510	0.580	0.8547
	Asa9	0.720	0.540	0.0002
	Asa16	0.300	0.000	0.0000
	Mean	0.470	0.310	
Lake Michigan	Asa2	-*	-*	-*
	Asa4	0.130	0.060	0.0323
	Asa8	0.080	0.080	1.0000
	Asa9	0.340	0.310	0.6003
	Asa16	0.340	0.000	0.0000
	Mean	0.220	0.110	

Population differentiation

Population differentiation was estimated using F_{ST} and R_{ST} values, which are summarized in table 8. F_{ST} values across Asa2, Asa8 and Asa9 show that there was moderate genetic differentiation among the populations. F_{ST} at Asa4 indicates little genetic differentiation. R_{ST} values for Asa9 shows greatest differentiation among populations compared to other loci. F_{ST} values for Asa4 and Asa16 indicates little genetic differentiation across populations. Both F_{ST} and R_{ST} for all populations across all loci show that there was significant population structuring among all populations.

Table 8: F_{ST} and R_{ST} across loci for all populations

Locus	F_{ST}	R_{ST}
Asa2	0.1409	-0.0109
Asa4	0.0018	0.0499
Asa8	0.0697	-0.0001
Asa9	0.1362	0.2701
Asa16	-0.0022	0.0338
All loci	0.1022	0.0699

0-0.05: little genetic differentiation
0.05-0.15: moderate differentiation
0.15-0.25: great differentiation
Above 0.25: very great genetic differentiation

The genetic differentiation between the Connecticut samples was assessed and each Connecticut population was compared with the more distant Lake Michigan population. Table 9 summarizes the details of these results. Since data were not available for Lake Michigan samples at the Asa2 locus, only the Connecticut samples were compared. Bride Brook and Roaring Brook populations displayed significant differentiation at this locus for F_{ST}, whereas there was little or no differentiation at all other loci between those two populations. Comparison between the two Connecticut populations and Lake Michigan population showed variable differentiation between loci. Lake Michigan population seems to be genetically distinct when

compared to the Connecticut samples at all loci except Asa16. The overall analysis for all loci for both F_{ST} and R_{ST} suggests that all the Connecticut populations may be substantially differentiated from Lake Michigan population and that the Connecticut samples are moderately differentiated from each other.

Table 9: Estimation of F_{ST} and R_{ST}

Locus	Population pair	F_{ST}	R_{ST}
Asa2	BB Vs RB	0.1409	-0.0109
	BB Vs LM	-	-
	RB Vs LM	-	-
Asa4	BB Vs RB	0.0023	-0.0001
	BB Vs LM	-0.0126	0.0863
	RB Vs LM	0.0226	0.0806
Asa8	BB Vs RB	0.0175	-0.0118
	BB Vs LM	0.0639	0.0371
	RB Vs LM	0.1544	-0.0061
Asa9	BB Vs RB	0.0077	0.0125
	BB Vs LM	0.2522	0.3876
	RB Vs LM	0.2843	0.4963
Asa16	BB Vs RB	0.0144	0.0678
	BB Vs LM	-0.0163	-0.0260
	RB Vs LM	-0.0244	-0.0192
All loci	BB Vs RB	0.0580	-0.0036
	BB Vs LM	0.1523	0.2652
	RB Vs LM	0.1847	0.3669

0-0.05: little genetic differentiation
0.05-0.15: moderate differentiation
0.15-0.25: great differentiation
Above 0.25: very great genetic differentiation

Discussion

The microsatellite loci used for alewife populations had enough genetic variation, with around 42 alleles across 6 loci. The loci were polymorphic and the genotypic distribution frequencies for all pairs of populations across all loci were significantly different, suggesting genetic structuring among populations (data not shown). Heterozygote deficiencies observed were different for different loci. Heterozygote deficiency can be interpreted as increase in homozygotes which might be a result of increased inbreeding. The significant heterozygote deficiency might reflect the fact that there is restricted gene flow between these anadromous populations. Although heterozygosities were low, there was enough variation present to examine any potential genetic differences among sample sites.

The fixation indices across all loci also indicate that the populations were significantly differentiated and that the two Connecticut samples differed from the landlocked Lake Michigan samples more distinctly than they did from each other. Lake Michigan samples are landlocked and more distant geographically to either of the Connecticut samples which are anadromous populations. Some gene flow due to straying might be expected among river populations, but not in landlocked populations. Hence the genetic differentiation between Connecticut populations is very subtle compared to Lake Michigan populations. Absence of products for Asa2 under the same PCR conditions for the landlocked Lake Michigan samples might further support the fact that this population is genetically distinct from the two anadromous populations. The F_{ST} and R_{ST} estimates suggest that differentiation might increase with geographic distance, though the number of populations under consideration in this study is too small to draw any such conclusions.

CHAPTER 3

Implications of the study on management of fish populations

The aim of this study was to estimate genetic differentiation in alewife populations using microsatellite loci. Analysis of neutral molecular markers has proven to be a robust method for identifying reproductive isolation among fish populations and aiding in their conservation and management (King et al, 2001). Alewife are anadromous fish which have high reproductive fidelity. We therefore predict that alewife populations should show genetic structuring among different natal sites across their geographical distribution. The microsatellite analysis suggests that identifiable population structuring exists among the sampled alewife populations.

Microsatellite data was used to estimate the population genetic statistics. Deviations from Hardy-Weinberg equilibrium and fixation indices like F_{ST} and R_{ST} were used to estimate population structuring. The significant heterozygote deficiency observed across the loci, which can be interpreted as increased inbreeding, reflects the fact that there might be restricted gene flow between these populations. This is consistent with the expectations since inter-breeding would not be possible between anadromous populations returning to their natal sites to spawn.

The fixation indices across all loci also indicate that the populations were genetically differentiated and that the two Connecticut samples differed from the Lake Michigan samples more distinctly than they did from each other. This is also consistent with the expectations since Lake Michigan samples are more distant geographically to either of the Connecticut samples and are landlocked in contrast to the anadromous river populations of Connecticut. Lake Michigan populations hence cannot interbreed with the Connecticut populations. The relationship between genetic distance and the geographic distance suggests that gene flow is restricted and follows a pattern of isolation by distance. Lack of homing or high rates of straying among populations may prevent the formation of population structure (McLean and Taylor, 2001). Some gene flow due to straying might be expected among river populations, but not in landlocked populations. This could explain why the Connecticut populations have very little differentiation compared to Lake

Michigan populations. Absence of products for Asa2 locus under the same PCR conditions for the landlocked Lake Michigan samples might further support the fact that this population is genetically distinct from the two anadromous populations. Thus the F_{ST} and R_{ST} estimates suggest that the populations are genetically distinct and that differentiation might increase with geographic distance.

Genetic structuring in anadromous populations depends on the fidelity rates of return to the natal streams. If there is considerable straying and migration between populations which translates to high rates of gene flow, it might result in swamping the effects of natural selection, homogenizing the populations. Hence the population structure depends on philopatry, and gene flow versus natural selection in these populations. Subtle genetic differences in the Connecticut populations suggest that there is straying and hence gene flow between populations, but observable local adaptations and adaptive divergence between these populations shows that selection is strong enough to act on these populations even in the presence of gene flow.

Behavioral limits to dispersal could restrict gene flow and allow for the accumulation of not only genetic differences but also morphological and physiological differences. Such differences were observed between the Roaring Brook and Bride Brook alewife populations in southern New England (Connecticut). Physiological differences, with respect to the location at which they spawn, were observed in these populations. Female and male alewife running up Bride Brook have relatively larger gonads than alewife running into Roaring Brook. There are also age differences among fish running in the two streams; the Bride Brook population's modal age is 3 yrs and the Roaring brook population's modal age is 4 yrs (E.Schultz, personal communication). These adaptive differences coupled with the observed genetic differences should provide considerable information to conservation biologists and assist in decision making for restocking programs.

Comparison of data analysis with other anadromous fish

Microsatellite analysis in shad by Waters et al (2000), for which these loci were actually developed, showed little differentiation among the populations sampled. The expected and

observed heterozygosities were almost similar across all loci for all the Atlantic populations in shad, where as observed heterozygosities were less compared to expected heterozygosities for almost all loci in alewife populations. Fixation indices were also lower in the Atlantic populations of shad with overall F_{ST} of 0.0063 while F_{ST} across all loci for alewife is 0.1022. Similarly, R_{ST} for shad over all loci was 0.0123 while it is 0.0699 for the alewife samples. Analysis of microsatellite loci for other anadromous fish showed F_{ST} estimates to range from 0.013 to 0.045 and R_{ST} to range around 0.036 and these fish populations were managed as distinct stocks (Lundy et al, 1999; McLean et al, 1999). The F_{ST} values in our analysis did fall in the above mentioned range for the Connecticut samples but the relatively high values of fixation indices over all loci for all alewife populations might be due to greater genetic differentiation in Lake Michigan samples compared to the Connecticut samples.

Limitations of microsatellite markers

Though microsatellites prove to be effective molecular markers, they have some limitations too. In most cases, microsatellites are shown to give poor estimates of gene flow between populations (Templeton, 1998). Another important limitation is size homoplasy which alters the estimation of population divergence when using statistics based on mutation models (Wilson et al, 2004). Rousset (1996) has shown that there is no effect of homoplasy on the parameters F_{IS} and R_{IS} and no simple effect on parameters F_{ST} and R_{ST} (Estoup et al, 2002).

However, since the effect of homoplasy is said to be reduced under larger scale (Jarne and Lagoda, 1996), further analyses with greater number of samples at each site would be necessary in order to get more reliable results. Furthermore, other sampling sites in Connecticut and Massachusetts should also be considered in the study to reveal more accurate patterns of genetic structuring of alewife populations in southern New England.

Management implications

Maintaining genetic integrity of natural stocks should likely be the priority of any management (Wilson et al, 2004). A precautionary approach to management usually aims at conserving inter- and intra-population variation in order to maximize the economic value of the resource. However, a fully precautionary approach should accommodate both local adaptation and population structuring. When transfer among locations is necessary, matched stocks from similar environments in nearby catchments should be used (Youngson et al, 2003).

It is also important to look at morphological, physiological and behavioral aspects of the native and exotic populations during stocking procedures. Mixing of divergent populations might effect their adaptations and hence their survival. There might also be differences in the fidelity rates or the levels of straying which determine the extent of genetic variation in the populations.

Subtle but significant genetic differentiation among the alewife populations suggest that migration between fishing grounds could be restricted to some extent and is therefore relevant to fisheries management. Since alewife populations in Connecticut show not only genetic differences but also differences in physiological characteristics, they should be treated with care while taking any restocking measures. It might be advisable to treat them as different stocks.

Conclusion

Though our data from the microsatellite loci indicate that there is significant differentiation among the alewife populations, further studies involving more sampling sites are required to make informed decisions regarding optimal restocking strategies. Nevertheless, our studies indicate that the microsatellite loci developed for shad can be used for alewife and that they are useful for detecting genetic structuring among these populations, the knowledge of which can be applied in conservation and restocking programs.

Future studies

As mentioned, further analyses which include more sampling sites and may be greater number of samples at each site would be necessary to reveal a more accurate pattern of genetic

structuring in alewife populations in southern New England. Also, comparison with mitochondrial DNA markers could reveal more appropriate information regarding population structuring and also gene flow among these populations. Patterns of genetic differentiation should also be consistent through time. Hence, more spatial and temporal studies on population structuring in alewife populations should be considered before planning any management programs for their restocking.

REFERENCES

Adams R I, Brown K M, and Hamilton M B (2004).The impact of microsatellite electromorph size homoplasy on multilocus population structure estimates in a tropical tree (*Corythophora alta*) and an anadromous fish (*Morone saxatilis*). *Molecular Ecology 13*: 2579-2588.

Aprahamian M W, Smith K M, McGinnity P, McKelvey S and Taylor J (2003). Restocking of Salmonids-opportunities and limitations. *Fisheries Research 62*: 211-227.

Balloux F and Lugon-Moulin N (2001). The estimation of population differentiation with microsatellite markers. *Molecular Ecology 11*: 155-165.

Beheregaray L B and Sunnucks P (2000). Microsatellite loci isolated from *Odontesthes argentinensis* and the *O. perugiae* species group and their use in other South American silverside fish. *Molecular Ecology 9 (5)*: 629-631.

Bouza C, Arias J, Castro J, Sanchez L and Martinez P (1999). Genetic structure of brown trout, *Salmo trutta* L., at the southern limit of the distribution range of the anadromous form. *Molecular Ecology 8*: 1991-2001.

Brown B L, Smouse P E, Epifanio J M and Kobak C J (1999). Mitochondrial DNA mixed-stock analysis of American Shad: Coastal harvests are dynamic and variable. *Transactions of the American Fisheries Society 128 (6)*: 977-994.

Brown B L, Gunter T P, Waters J M and Epifanio J M (2000). Evaluating genetic diversity associated with propagation-assisted restoration of American Shad. *Conservation Biology 14 (1)*:291-303.

Brown K M, Baltazar G A, Weinstein B N and Hamilton M B (2003). Isolation and characterization of nuclear microsatellite loci in the anadromous fish *Morone saxatilis. Molecular Ecology Notes 3*: 414-416.

Burridge C P and Smolenski A J (2000). Microsatellite loci from the marine fish *Nemadactylus macropterus* (Perciformes: Cheilodactylidae). *Molecular Ecology 9 (8)*: 1180-1181.

Corkett C J (2005). The Pew report on US fishery councils: a critique from the open society. *Marine Policy 29:* 247-253.

D'Amato M E, Lunt D H, Carvalho G R (1999). Microsatellite markers for the hake *Macruronus magellanicus* amplify other gadoid fish. *Molecular Ecology 8 (6):* 1086-1088.

Daniels R A (2001). Untested assumptions: The role of canals in the dispersal of sea lamprey, alewife, and other fishes in the eastern United States. *Environmental Biology Of Fishes 60 (4)*: 309-329.

deGroot S J (2002). A review of the past and present status of anadromous fish species in the Netherlands: is restocking the Rhine feasible? *Hydrobiologia 478*: 205-218.

Estoup A, Jarne P and Cornuet J (2002). Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology 11*: 1591-1604.

Gulseth O A, Nilssen K J, Iversen M, Finstad B (2001). Seawater tolerance in first-time migrants of anadromous Arctic charr (*Salvelinus alpinus*). *Polar Biology 24*: 270-275.

Gunderina L I (2003). DNA polymorphism of *Drosophila* genes and factors determining it. *Russian Journal of Genetics 39 (7)*: 737-747.

Hendry A P (2001). Adaptive divergence and the evolution of reproductive isolation in the wild: an empirical demonstration using introduced sockeye salmon. *Genetica 112-113*: 515-534.

Hinten G, Harriss F, Rossetto M, and Braverstock P R (2003). Genetic variation and island biogeography: Microsatellite and mitochondrial DNA variation in island populations of the Australian bush rat, *Rattus fuscipes greyii. Conservation Genetics 4*: 759-778.

Hoffmann A A and Merila J (1999). Heritable variation and evolution under favourable and unfavourable conditions. *Trends in Ecology and Evolution 14 (3)*: 96-101.

Jarne P and Lagoda P (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution 11(10)*: 424-429.

Jorgensen E H and Arnesen A M (2002). Seasonal changes in osmotic and ionic regulation in Arctic charr, *Salvelinus alpinus,* from a high- and a sub-arctic anadromous population. *Environmental Biology of Fishes 64*: 185-193.

King T L, Kalinowski S T, Schill W B, Spidle A P and Lubinski B A (2001). Population structure of Atlantic salmon (*Salmo salar* L.): a range-wide perspective from microsatellite DNA variation. *Molecular Ecology 10*: 807-821.

Kosa J T and Mather M E (2001). Processes contributing to variability in regional patterns of juvenile river herring abundance across small coastal systems. *Transaction of the American Fisheries Society 130*: 600-619.

Krayushkina L S, Gerasimov A A, and Smirnov A V (2001). Hypoosmotic regulation in anadromous marine sturgeon, with special reference to the structure and function of their kidneys and gill chloride cells. *Doklady Biological Sciences 378*: 210-212.

Lenormand T (2002). Gene flow and the limits to natural selection. *Trends in Ecology and Evolution 17 (4)*: 183-189.

Limburg K E (1998). Anomalous migrations of anadromous herrings revealed with natural chemical tracers. *Canadian Journal of Fisheries and Aquatic Sciences 55* (2): 431-437.

Luikart G and England P R (1999). Statistical analysis of microsatellite DNA data. *Trends in Ecology and Evolution 14 (7)*: 253-255

Lundy C J, Moran P, Rico C, Milner R S and Hewitt G M (1999). Macrogeographical population differentiation in oceanic environments: a case study of European hake (*Merluccius merluccius*), a commercially important fish. *Molecular Ecology 8*: 1889-1898.

Mackay T F (1980). Genetic variance, fitness, and homeostasis in varying environments: an experimental check of the theory. *Evolution 34 (6)*: 1219-1222.

Madeira M J, Gomez-Moliner B J and Barbe A M (2005). Genetic introgression on freshwater fish populations caused by restocking programmes. *Biological Invasions 7*: 117-125.

Madenjian C P, O'Connor D V, Chernyak S M, Rediske R R, and O'Keefe J P (2004). Evaluation of Chinook salmon (*Oncorhynchus tshawytscha*) bioenergetics model. *Canadian Journal of Fisheries and Aquatic Sciences 61 (4)*: 627-635.

Marzano F N, Corradi N, Papa R, Tagliavini J and Gandolfi G (2003). Molecular evidence for introgression and loss of genetic variability in *Salmo (trutta) macrostigma* as a result of massive restocking of Apennine populations (Northern and Central Italy). *Environmental Biology of Fishes 68*: 349-356.

Matocq M D and Lacey E A (2004). Philopatry, kin clusters and genetic relatedness in a population of woodrats (*Neotoma macrotis*). *Behavioral Ecology 15(4):* 647-653.

McDowall R M (1997). The evolution of diadromy in fishes (revisited) and its place in phylogenetic analysis. *Reviews in Fish Biology and Fisheries 7*: 443-462.

McKay J K and Latta R G (2002). Adaptive population divergence: markers, QTL and traits. *Trends in Ecology and Evolution 17 (6)*: 285-291.

McLean J E, Hay D E and Taylor E B (1999). Marine population structure in an anadromous fish: life-history influences patterns of mitochondrial DNA variation in the eulachon, *Thaleichthys pacificus. Molecular Ecology 8*: S143-S158.

McLean J E and Taylor E B (2001). Resolution of population structure in a species with high gene flow: microsatellite variation in the eulachon (Osmeridae: *Thaleichthys pacificus*). *Marine Biology 139*: 411-420.

Meyers L A and Bull J J (2002). Fighting change with change: adaptive variation in an uncertain world. *Trends in Ecology and Evolution 17 (12)*: 551-557.

Miller L M, Kallemeyn L and Senanan W (2001). Spawning-site and natal-site fidelity by northern pike in a large lake: mark-recapture and genetic evidence. *Transactions of the American Fisheries Society 130*: 307-316.

Mitchell-Olds T (1995). The molecular basis of quantitative genetic variation in natural populations. *Trends in Ecology and Evolution 10 (8)*:324-328.

Moring J R and Mink L H (2002). Anadromous alewives, *Alosa pseudoharengus,* as prey for white perch, *Morone americana. Hydrobiologia 479*: 125-130.

Munroe T A (2002), *Herrings. Family Clupeidae.* Pp 111-160 in: Bigelow and Schroeder's Fishes of the Gulf of Maine. Ed. B. B. Collette and G. Klein-MacPhee. Washington, D.C: Smithsonian Press

Nielsen J L, Crow K D and Fountain M C (1998). Microsatellite diversity and conservation of a relic trout population: McCloud River redband trout. *Molecular Ecology 8*: S129-S142.

O'Connell M and Wright J M (1997). Microsatellite DNA in fishes. *Reviews in Fish Biology and Fisheries 7*: 331-363.

Paterson S, Piertney S B, Knox D, Gilbey J, and Verspoor E (2004). Characterization and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon (*Salmo salar* L.) microsatellites. *Molecular Ecology Notes 4*: 160-162.

Prignon C, Micha J C, Rimbaud G and Philippart J C (1999). Rehabilitation efforts for Atlantic salmon in the Meuse basin area: Synthesis 1983-1998. *Hydrobiologia 410*: 69-77.

Quattro J M, Greig T W, Coykendall D K, Bowen B W and Baldwin J D (2002). Genetic issues in aquatic species management: the shortnose sturgeon (*Acipenser brevirostrum*) in the southeastern United States. *Conservation Genetics 3*: 155-166

Reilly A, Elliott N G, Grewe P M, Clabby C, Powell R, Ward R D (1999). Genetic differentiation between Tasmanian cultured Atlantic salmon (*Salmo salar* L.) and their ancestral Canadian population: comparison of microsatellite DNA and allozyme and mitochondrial DNA variation. *Aquaculture 173*: 459-469.

Raymond M and Rousset F (1995). GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity 86:* 248-249.

Rousset F (1996). Equilibrium values of measures of population subdivision for stepwise mutation process. *Genetics 142*: 1357-1362.

Sambrook J, Fritsch E, Maniates T (1989). Molecular cloning, a laboratory Manual 2nd edition, Cold Spring Harbor Laboratory, NY.

Schlotterer C and Harr B (2001). Microsatellite instability. *Encyclopedia of Life Sciences*: 1-4

Scholz N L, Truelove N K, French B L, Berejikian B A, Quinn T P, Casillas E, and Collier T K (2000). Diazinon disrupts antipredator and homing behaviors in chinook salmon (*Oncorhynchus tshawytscha*). *Canadian Journal of Fisheries and Aquatic Sciences 57*: 1911-1918.

Slatkin M (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics 139*: 457-462.

Smith T B, Schneider C J and Holder K (2001). Refugial isolation versus ecological gradients. *Testing alternative mechanisms of evolutionary divergence in four rainforest vertebrates. Genetica 112-113*: 383-398.

Spieth P T (1974). Gene flow and genetic differentiation. *Genetics 78:* 961-965.

Steiner U K and Gaston A J (2005). Reproductive consequences of natal dispersal in a highly philopatric seabird. *Behavioral Ecology 16*: 634-639.

Storfer A (1999). Gene flow and local adaptation in a sunfish – salamander system. *Behavioral Ecology and Sociobiology 46:* 273-279

Sunnucks P (2000). Efficient genetic markers for population biology. *Trends in Ecology and Evolution 15 (5)*: 199-203.

Taylor E B (1999). Species pairs of north temperate freshwater fishes: Evolution, taxonomy, and conservation. *Reviews in Fish Biology and Fisheries 9*: 299-324.

Templeton A R (1998). Nested clade analysis of phylogeographic data: testing hypothesis about gene flow and population history. *Molecular Ecology 7*:381-397.

Tessier N, Bernatchez L and Wright J M (1997). Population structure and impact of supportive breeding inferred from mitochondrial and microsatellite DNA analyses in land-locked Atlantic salmon *Salmo salar* L *Molecular Ecology 6:* 735-750

Tregenza T (2002). Divergence and reproductive isolation in the early stages of speciation. *Genetica 116:* 291-300.

Uthicke S and Purcell S (2004). Preservation of genetic diversity in restocking of the sea cucumber. *Holothuria scabra* investigated by allozyme electrophoresis. *Canadian Journal of Fisheries and Aquatic Sciences 61:* 519-528.

Waters J M, Esa Y B, Wallis G P (1999). Characterization of microsatellite loci from a New Zealand freshwater fish (*Galaxias vulgaris*) and their potential for analysis of hybridization in Galaxiidae. *Molecular Ecology 8 (6):* 1080-1082

Waters J M, Epifanio J M, Gunter T, and Brown B L (2000). Homing behaviour facilitates subtle genetic differentiation among river populations of *Alosa sapidissima*: microsatellites and mtDNA. *Journal of Fish Biology 56*: 622-636.

Weir B S and Cockerham C C (1984). Estimating F statistics for the analysis of population structure. *Evolution 38 (6)*: 1358-1370.

Wilson A J, Gislason D, Skulason S, Snorrason S S, Adams C E, Alexander G, Danzmann R G, and Ferguson M M (2004). Population genetic structure of Arctic Charr, *Salvelinus alpinus* from northwest Europe on large and small spatial scales. *Molecular Ecology 13*: 1129-1142.

Wood C C and Foote C J (1996). Evidence for sympatric genetic divergence of anadromous and non-anadromous morphs of sockeye salmon (*Oncorhynchus nerka*). *Evolution 50 (3):* 1265-1279.

Wright S (1965). The interpretation of population structure by F-Statistics with special regard to system of mating. *Evolution 19 (3):* 395-420.

Youngson A F, Jordan W C, Verspoor E, McGinnity P, Cross T and Ferguson A (2003). Management of salmonid fisheries in the British Isles: towards a practical approach based on population genetics. *Fisheries Research 62:* 193-209.

Zoller S, Lutzoni F and Scheidegger C (1999). Genetic variation within and among populations of the threatened liken *lobaria pulmonaira* in Switzerland and implications for its conversation. *Molecular Ecology 8:* 2049-2059

APPENDIX A: Unedited and edited genotyping data for the 5 microsatellite loci in alewife.

Scoring alleles: The smallest allele was assigned the number 1 and other alleles were assigned subsequent numbers based on the kind of repeat. For trinucleotide repeats alleles scored as 3 bases larger than the smallest allele were designated as allele 2 and and alleles scored as 6 bases larger than the smallest allele were designated as allele 3 and so on. Tetranucleotide repeats were also scored in a similar fashion but taking a difference of 4 bases as a new allele.

DNW= did not work

Asa2 Locus								
Sample	**Unedited Allele 1**	**Edited Allele 1**	**Unedited Size 1**	**Edited Size 1**	**Unedited Allele 2**	**Edited Allele 2**	**Unedited Size 2**	**Edited Size 2**
BB-1	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-2	9	9	106.89	106.89	9	9	106.89	106.89
BB-3	9	9	106.5	106.5	9	9	106.5	106.5
BB-4	5	5	97.84	97.84	8	8	103.88	103.88
BB-5	5	5	97.67	97.67	5	5	97.67	97.67
BB-6	0	0	DNW	DNW	0	0	DNW	DNW
BB-7	0	0	DNW	DNW	0	0	DNW	DNW
BB-8	6	6	100.9	100.9	6	6	100.9	100.9
BB-9	0	0	DNW	DNW	0	0	DNW	DNW
BB-10	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-11	8	8	103.77	103.77	8	8	103.77	103.77
BB-12	8	8	103.84	103.84	8	8	103.84	103.84
BB-13	8	8	103.86	103.86	8	8	103.86	103.86
BB-14	8	8	103.87	103.87	8	8	103.87	103.87
BB-15	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-16	0	0	DNW	DNW	0	0	DNW	DNW
BB-17	0	0	DNW	DNW	0	0	DNW	DNW
BB-18	0	0	DNW	DNW	0	0	DNW	DNW
BB-19	19	19	125.47	125.47	19	19	125.47	125.47
BB-20	19	19	125.18	125.18	19	19	125.18	125.18
BB-21	0	0	DNW	DNW	0	0	DNW	DNW
BB-22	8	8	104.46	103.46	8	8	104.46	103.46
BB-23	5	5	97.7	97.7	5	5	97.7	97.7
BB-24	1	1	79.61	79.61	1	1	79.61	79.61
BB-25	8	8	103.84	103.84	8	8	103.84	103.84
BB-26	6	6	100.78	100.78	6	6	100.78	100.78
BB-27	8	8	103.87	103.87	8	8	103.87	103.87
BB-28	0	0	DNW	DNW	0	0	DNW	DNW
BB-29	8	8	103.59	103.59	8	8	103.59	103.59
BB-30	8	8	103.72	103.72	8	8	103.72	103.72
bb-1	0	0	DNW	DNW	0	0	DNW	DNW
bb-2	8	8	104.08	103.08	8	8	104.08	103.08

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
bb-3	0	0	DNW	DNW	0	0	DNW	DNW
bb-4	8	8	104.34	103.34	8	8	104.34	103.34
bb-5	0	0	DNW	DNW	0	0	DNW	DNW
bb-6	0	0	DNW	DNW	0	0	DNW	DNW
bb-7	0	0	DNW	DNW	0	0	DNW	DNW
bb-8	0	0	DNW	DNW	0	0	DNW	DNW
bb-9	8	8	104.14	103.14	8	8	104.14	103.14
bb-10	0	0	DNW	DNW	0	0	DNW	DNW
bb-11	0	0	DNW	DNW	0	0	DNW	DNW
bb-12	0	0	DNW	DNW	0	0	DNW	DNW
bb-13	0	0	DNW	DNW	0	0	DNW	DNW
bb-14	0	0	DNW	DNW	0	0	DNW	DNW
bb-15	0	0	DNW	DNW	0	0	DNW	DNW
bb-16	0	0	DNW	DNW	0	0	DNW	DNW
bb-17	0	0	DNW	DNW	0	0	DNW	DNW
bb-18	0	0	DNW	DNW	0	0	DNW	DNW
bb-19	0	0	DNW	DNW	0	0	DNW	DNW
bb-20	0	0	DNW	DNW	0	0	DNW	DNW
bb-21	0	0	DNW	DNW	0	0	DNW	DNW
bb-22	0	0	DNW	DNW	0	0	DNW	DNW
bb-23	0	0	DNW	DNW	0	0	DNW	DNW
bb-24	0	0	DNW	DNW	0	0	DNW	DNW
bb-25	0	0	DNW	DNW	0	0	DNW	DNW
bb-26	0	0	DNW	DNW	0	0	DNW	DNW
RB-1	8	8	103.97	103.97	8	8	103.97	103.97
RB-2	0	0	DNW	DNW	0	0	DNW	DNW
RB-3	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-4			DNW	DNW			DNW	DNW
RB-5	8	8	103.99	103.99	8	8	103.99	103.99
RB-6	0	0	DNW	DNW	0	0	DNW	DNW
RB-7	0	0	DNW	DNW	0	0	DNW	DNW
RB-8	0	0	DNW	DNW	0	0	DNW	DNW
RB-9	0	0	DNW	DNW	0	0	DNW	DNW
RB-10	7	6	101.06	100.06	10	9	107.17	106.17
RB-11	11	11	110.04	109.04	11	11	110.04	109.04
RB-12	8	8	104.14	103.14	8	8	104.14	103.14
RB-13	0	0	DNW	DNW	0	0	DNW	DNW
RB-14	0	0	DNW	DNW	0	0	DNW	DNW
RB-15	0	0	DNW	DNW	0	0	DNW	DNW
RB-16	11	11	110.22	109.22	11	11	110.22	109.22
RB-17	0	0	DNW	DNW	0	0	DNW	DNW
RB-18	0	0	DNW	DNW	0	0	DNW	DNW
RB-19	8	8	104.12	103.12	8	8	104.12	103.12
RB-20	0	0	DNW	DNW	0	0	DNW	DNW
RB-21	11	11	110.15	109.15	13	13	113.09	113.09

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
RB-23	0	0	DNW	DNW	0	0	DNW	DNW
RB-24	0	0	DNW	DNW	0	0	DNW	DNW
RB-25	0	0	DNW	DNW	0	0	DNW	DNW
RB-26	8	8	104.47	103.47	10	11	110.47	109.47
RB-27	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-28	13	13	113.29	113.29	13	13	113.29	113.29
RB-29	11	11	109.92	109.92	11	11	109.92	109.92
RB-30	0	0	DNW	DNW	0	0	DNW	DNW
RB-31	11	11	109.78	109.78	11	11	109.78	109.78
RB-32	11	11	110.09	109.09	11	11	110.09	109.09
RB-33	22	22	132.92	132.92	30	30	148.79	148.79
RB-34	12	11	111.47	109.47	12	11	111.47	109.47
RB-35	13	13	113.22	113.22	13	13	113.22	113.22
RB-36	13	13	113.04	113.04	13	13	113.04	113.04
RB-37	0	0	DNW	DNW	0	0	DNW	DNW
RB-38	0	0	DNW	DNW	0	0	DNW	DNW
RB-39	0	0	DNW	DNW	0	0	DNW	DNW
RB-40	0	0	DNW	DNW	0	0	DNW	DNW
RB-41	0	0	DNW	DNW	0	0	DNW	DNW
RB-42	0	0	DNW	DNW	0	0	DNW	DNW
RB-43	3	3	85.71	85.71	3	3	85.71	85.71
RB-44	2	2	83.1	83.1	4	4	87.46	87.46
RB-45	2	2	83.14	83.14	4	4	87.47	87.47
RB-46	8	8	103.83	103.83	11	11	109.85	109.85
RB-47	2	2	83.21	83.21	4	4	87.55	87.55
RB-48	0	0	DNW	DNW	0	0	DNW	DNW
RB-49	0	0	DNW	DNW	0	0	DNW	DNW
RB-50	0	0	DNW	DNW	0	0	DNW	DNW
RB-51	0	0	DNW	DNW	0	0	DNW	DNW
RB-52	0	0	DNW	DNW	0	0	DNW	DNW
RB-53	2	2	83.16	82.16	4	4	87.47	87.47
RB-54	0	0	DNW	DNW	0	0	DNW	DNW
RB-55	0	0	DNW	DNW	0	0	DNW	DNW
RB-56	0	0	DNW	DNW	0	0	DNW	DNW
RB-57	0	0	DNW	DNW	0	0	DNW	DNW

Locus Asa4								
Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
BB-1	8	8	117.16	117.16	9	8	120.01	117.01
BB-2	8	8	117.04	117.04	9	8	119.93	117.93
BB-3	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-4	9	8	120.01	117.01	10	10	122.87	122.87
BB-5	8	8	117.16	117.16	9	8	120.01	117.01
BB-6	9	8	120.49	117.49	13	13	128.79	128.79
BB-7	8	8	117.3	117.3	9	8	120.11	117.11
BB-8	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-9	8	8	117.09	117.09	9	8	119.92	117.92
BB-10	8	8	117.11	117.11	9	8	119.98	117.98
BB-11	8	8	117.04	117.04	9	8	119.89	117.89
BB-12	8	8	117.29	117.29	8	8	117.29	117.29
BB-13	8	8	117.07	117.07	9	8	119.9	117.9
BB-14	8	8	117.07	117.07	9	8	119.9	117.9
BB-15	8	8	117.2	117.2	9	8	120.01	117.01
BB-16	8	8	117.04	117.04	9	8	119.97	117.97
BB-17	8	8	117.07	117.07	9	8	119.9	117.9
BB-18	8	8	117.09	117.09	9	8	119.97	117.97
BB-19	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-20	7	8	116.94	116.94	9	8	119.82	117.82
BB-21	8	8	117.04	117.04	9	8	119.89	117.89
BB-22	8	8	117.07	117.07	9	8	119.9	117.9
BB-23	8	8	117.15	117.15	9	8	120.13	117.13
BB-24	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-25	0	0	DNW	DNW	0	0	DNW	DNW
BB-26	7	8	116.93	116.93	9	8	119.8	117.8
BB-27	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-28	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-29	8	8	117.08	117.08	9	8	119.94	117.94
BB-30			119.89	117.89	9	8	119.89	117.89
bb-1	8	8	117.07	117.07	9	8	119.9	117.9
bb-2	8	8	117.18	117.18	9	8	120.1	117.1
bb-3	8	8	117.27	117.27	9	8	120.15	117.15
bb-4	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
bb-5	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
bb-6	8	8	117.18	117.18	9	8	120.03	117.03
bb-7	8	8	117.15	117.15	9	8	119.98	117.98

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
bb-8	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
bb-9	8	8	117.15	117.15	9	8	119.98	117.98
bb-10	8	8	117.25	117.25	8	8	117.25	117.25
bb-11	8	8	117.17	117.17	8	8	117.17	117.17
bb-12	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
bb-13	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
bb-14	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
bb-15	9	8	120.05	117.05	9	8	120.05	117.05
bb-16	8	8	117.22	117.22	12	12	125.98	125.98
bb-17	0	0	DNW	DNW	0	0	DNW	DNW
bb-18	0	0	DNW	DNW	0	0	DNW	DNW
bb-19	0	0	DNW	DNW	0	0	DNW	DNW
bb-20	0	0	DNW	DNW	0	0	DNW	DNW
bb-21	0	0	DNW	DNW	0	0	DNW	DNW
bb-22	0	0	DNW	DNW	0	0	DNW	DNW
bb-23	0	0	DNW	DNW	0	0	DNW	DNW
bb-24	0	0	DNW	DNW	0	0	DNW	DNW
bb-25	0	0	DNW	DNW	0	0	DNW	DNW
bb-26	0	0	DNW	DNW	0	0	DNW	DNW
RB-1	8	8	117.16	117.16	9	8	119.96	117.96
RB-2	8	8	117.16	117.16	9	8	119.96	117.96
RB-3	8	8	117.21	117.21	9	8	120.04	117.04
RB-4	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-5	8	8	117.08	117.08	9	8	119.89	117.89
RB-6	8	8	117.21	117.21	9	8	120.15	117.15
RB-7	8	8	117.3	117.3	9	8	120.11	117.11
RB-8	8	8	117.38	117.38	9	8	120.18	117.18
RB-9	8	8	117.16	117.16	9	8	119.96	117.96
RB-10	8	8	117.16	117.16	9	8	119.96	117.96
RB-11	8	8	117.18	117.18	9	8	119.98	117.98
RB-12	8	8	117.3	117.3	9	8	120.11	117.11
RB-13	8	8	117.18	117.18	9	8	119.98	117.98
RB-14	8	8	117.17	117.17	9	8	120	117
RB-15	8	8	117.25	117.25	9	8	120.07	117.07
RB-16	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-17	8	8	117.3	117.3	9	8	120.1	117.1
RB-18	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-19	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
RB-20	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-21	8	8	117.2	117.2	9	8	120.01	117.01
RB-23	8	8	117.25	117.25	9	8	120.07	117.07
RB-24	8	8	117.21	117.21	9	8	120.04	117.04
RB-25	8	8	117.3	117.3	9	8	120.1	117.1
RB-26	8	8	117.28	117.28	9	8	120.03	117.03
RB-27	8	8	117.22	117.22	9	8	120.13	117.13
RB-28	8	8	117.43	117.43	9	8	120.19	117.19
RB-29	8	8	117.2	117.2	9	8	119.96	117.96
RB-30	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-31	8	8	117.08	117.08	9	8	119.89	117.89
RB-32	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-33	8	8	117.22	117.22	9	8	120.18	117.18
RB-34	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-35	8	8	117.28	117.28	9	8	120.03	117.03
RB-36	8	8	117.26	117.26	9	8	120.17	117.17
RB-37	8	8	117.34	117.34	9	8	120.09	117.09
RB-38	8	8	117.27	117.27	9	8	120.21	117.21
RB-39	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-40	8	8	117.14	117.14	9	8	119.95	117.95
RB-41	8	8	117.13	117.13	9	8	120.08	117.08
RB-42	9	8	119.96	117.96	13	13	128.67	128.67
RB-43	8	8	117.09	117.09	9	8	119.97	117.97
RB-44	8	8	117.04	117.04	9	8	119.89	117.89
RB-45	8	8	117.04	117.04	9	8	119.89	117.89
RB-46	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-47	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-48	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-49	8	8	117.21	117.21	9	8	119.99	117.99
RB-50	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-51	9	8	120.01	117.01	9	8	120.01	117.01
RB-52	8	8	117.07	117.07	9	8	119.9	117.9
RB-53	8	8	117.16	117.16	9	8	120.01	117.01
RB-54	8	8	117.28	117.28	9	8	119.93	117.93
RB-55	8	8	117.14	117.14	9	8	119.94	117.94
RB-56	8	8	117.08	117.08	8	8	117.08	117.08
RB-57	8	8	117.3	117.3	9	8	120.11	117.11

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
LM-1	8	8	117.04	117.04	9	8	119.89	117.89
LM-2	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
LM-3	8	8	117.14	117.14	9	8	119.99	117.99
LM-4	7	8	116.94	116.94	9	8	119.95	117.95
LM-5	1	1	103.58	103.58	3	3	108.77	108.77
LM-6	8	8	117.2	117.2	9	8	120.01	117.01
LM-7	8	8	117.21	117.21	9	8	120.04	117.04
LM-8	8	8	117.23	117.23	9	8	120.05	117.05
LM-9	8	8	117.22	117.22	9	8	120.02	117.02
LM-10	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
LM-11	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
LM-12	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
LM-13	8	8	117.23	117.23	9	8	120.06	117.06
LM-14	8	8	117.28	117.28	9	8	120.08	117.08
LM-15	8	8	117.08	117.08	9	8	119.89	117.89
LM-16	8	8	117.4	117.4	9	8	120.2	117.2
LM-17	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
LM-18	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
LM-19	8	8	117.23	117.23	9	8	120.22	117.22
LM-20	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
LM-21	8	8	117.25	117.25	9	8	120.07	117.07
LM-22	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
LM-23	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
LM-24	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
LM-25	8	8	117.22	117.22	9	8	120.02	117.02

Locus Asa 8								
Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
BB-1	7	6	134.41	132.41	9	8	138.68	136.68
BB-2	7	6	133.49	132.49	7	6	133.49	132.49
BB-3	6	6	132.87	132.87	6	6	132.87	132.87
BB-4	6	6	132.79	132.79	6	6	132.79	132.79
BB-5	10	10	139.46	140.46	10	10	139.46	140.46
BB-6	6	6	132.94	132.94	6	6	132.94	132.94
BB-7	6	6	132.85	132.85	6	6	132.85	132.85
BB-8	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-9	6	6	132.97	132.97	9	8	137.2	136.2
BB-10	6	6	132.77	132.77	6	6	132.77	132.77
BB-11	7	6	134.63	132.63	9	8	138.68	136.68
BB-12	7	6	134.52	132.52	9	8	138.68	136.68
BB-13	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-14	4	4	128.65	128.65	6	6	132.83	132.83
BB-15	6	6	132.74	132.74	6	6	132.74	132.74
BB-16	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-17	3	3	124.79	124.79	7	6	133.44	132.44
BB-18	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-19	6	6	132.76	132.76	6	6	132.76	132.76
BB-20	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-21	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-22	6	6	132.93	132.93	6	6	132.93	132.93
BB-23	5	4	129.01	128.01	7	6	133.11	132.11
BB-24	7	6	133.17	132.17	7	6	133.17	132.17
BB-25	7	6	133.05	132.05	9	8	137.39	136.39
BB-26	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-27	6	6	132.66	132.66	6	6	132.66	132.66
BB-28	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
BB-29	6	6	132.79	132.79	6	6	137.2	137.2
BB-30	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
bb-1	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
bb-2	7	6	133.1	132.1	7	6	133.1	132.1
bb-3	6	6	132.87	132.87	9	8	137.7	136.7
bb-4	8	8	136.26	136.26	8	8	136.26	136.26

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
bb-5	6	6	132.8	132.8	6	6	132.8	132.8
bb-6	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
bb-7	6	6	132.99	132.99	6	6	132.99	132.99
bb-8	7	6	133	132	7	6	133	132
bb-9	7	6	133.13	132.13	9	8	137.41	136.41
bb-10	5	4	129.09	128.09	7	6	133.16	132.16
bb-11	6	6	132.96	132.96	6	6	132.96	132.96
bb-12	6	6	132.61	132.61	6	6	132.61	132.61
bb-13	6	6	132.72	132.72	6	6	132.72	132.72
bb-14	6	6	132.82	132.82	6	6	132.82	132.82
bb-15	6	6	132.86	132.86	6	6	132.86	132.86
bb-16	7	6	133.07	132.07	7	6	133.07	132.07
bb-17	0	0	DNW	DNW	0	0	DNW	DNW
bb-18	0	0	DNW	DNW	0	0	DNW	DNW
bb-19	0	0	DNW	DNW	0	0	DNW	DNW
bb-20	0	0	DNW	DNW	0	0	DNW	DNW
bb-21	0	0	DNW	DNW	0	0	DNW	DNW
bb-22	0	0	DNW	DNW	0	0	DNW	DNW
bb-23	0	0	DNW	DNW	0	0	DNW	DNW
bb-24	0	0	DNW	DNW	0	0	DNW	DNW
bb-25	0	0	DNW	DNW	0	0	DNW	DNW
bb-26	0	0	DNW	DNW	0	0	DNW	DNW
RB-1	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-2	6	6	132.88	132.88	9	8	137.22	136.22
RB-3	4	4	128.89	128.89	6	6	132.88	132.88
RB-4	0	0	DNW	DNW	0	0	DNW	DNW
RB-5	7	6	133.04	132.04	7	6	133.04	132.04
RB-6	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-7	7	6	133.2	132.2	9	8	137.54	136.54
RB-8	7	6	133.39	132.39	7	6	133.39	132.39
RB-9	5	4	129.32	128.32	7	6	133.52	132.52
RB-10	7	6	133.2	132.2	9	8	137.58	136.58
RB-11	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-12	7	6	133.33	132.33	7	6	133.33	132.33
RB-13	7	6	133.66	132.66	9	8	137.58	136.58
RB-14	7	6	133.31	132.31	9	8	137.57	136.57
RB-15	7	6	133.45	132.45	7	6	133.45	132.45
RB-16	7	6	133.52	132.52	9	8	137.74	136.74
RB-17	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
RB-18	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-19	6	6	132.91	132.91	6	6	132.91	132.91
RB-20	7	6	133.19	132.19	9	8	137.38	136.38
RB-21	7	6	133.22	132.22	9	8	137.39	136.39
RB-23	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-24	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-25	2	2	120.39	120.39	7	6	133.1	132.1
RB-26	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-27	7	6	133.28	132.28	7	6	133.28	132.28
RB-28	7	6	133.3	132.3	7	6	133.3	132.3
RB-29	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-30	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-31	7	6	133.19	132.19	9	8	137.74	136.74
RB-32	7	6	133.16	132.16	7	6	133.16	132.16
RB-33	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-34	7	6	133.04	132.04	7	6	133.04	132.04
RB-35	2	2	120.75	120.75	6	6	133.48	132.48
RB-36	7	6	133.03	132.03	7	6	133.03	132.03
RB-37	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-38	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-39	7	6	133.55	132.55	7	6	133.55	132.55
RB-40	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-41	7	6	133.08	132.08	9	8	137.23	136.23
RB-42	9	8	138.52	136.52	9	8	138.52	136.52
RB-43	6	6	132.74	132.74	6	6	132.74	132.74
RB-44	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-45	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-46	6	6	132.8	132.8	6	6	132.8	132.8
RB-47	4	4	128.96	128.96	7	6	133.32	132.32
RB-48	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-49	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-50	6	6	132.79	132.79	9	8	137.2	136.2
RB-51	6	6	132.74	132.74	6	6	132.74	132.74
RB-52	4	4	128.8	128.8	6	6	132.89	132.89

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
RB-53	2	2	120.41	120.41	5	4	129.23	128.23
RB-54	6	6	132.85	132.85	9	8	137.21	136.21
RB-55	7	6	134.65	132.65	9	8	138.68	136.68
RB-56	2	2	120.35	120.35	7	6	133.21	132.21
RB-57	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
LM-1	7	6	133.13	132.13	7	6	133.13	132.13
LM-2	7	6	133.05	132.05	7	6	133.05	132.05
LM-3	7	6	133.08	132.08	7	6	133.08	132.08
LM-4	7	6	133.08	132.08	7	6	133.08	132.08
LM-5	7	6	133.07	132.07	7	6	133.07	132.07
LM-6	7	6	133.12	132.12	7	6	133.12	132.12
LM-7	6	6	132.84	132.84	6	6	132.84	132.84
LM-8	5	4	129.78	128.78	7	6	133.26	132.26
LM-9	7	6	133.15	132.15	7	6	133.15	132.15
LM-10	6	6	132.89	132.89	6	6	132.89	132.89
LM-11	6	6	132.78	132.78	6	6	132.78	132.78
LM-12	6	6	132.93	132.93	6	6	132.93	132.93
LM-13	7	6	133.09	132.09	7	6	133.09	132.09
LM-14	6	6	132.95	132.95	6	6	132.95	132.95
LM-15	7	6	133.02	132.02	7	6	133.02	132.02
LM-16	1	1	116.81	116.81	7	6	133.31	132.31
LM-17	7	6	133.04	132.04	7	6	133.04	132.04
LM-18	6	6	132.87	132.87	6	6	132.87	132.87
LM-19	7	6	133.12	132.12	7	6	133.12	132.12
LM-20	7	6	133.13	132.13	7	6	133.13	132.13
LM-21	0	0	DNW	DNW	0	0	DNW	DNW
LM-22	7	6	133.19	132.19	7	6	133.19	132.19
LM-23	7	6	133.21	132.21	7	6	133.21	132.21
LM-24	7	6	133.37	132.37	7	6	133.37	132.37
LM-25	6	6	132.87	132.87	6	6	132.87	132.87

Locus Asa 9								
Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
BB-1	6	6	186.38	186.38	10	10	198.48	198.48
BB-2	7	7	190.32	190.32	7	7	190.32	190.32
BB-3	8	8	194.3	194.3	8	8	194.3	194.3
BB-4	3	3	178.33	178.33	8	8	194.51	194.51
BB-5	8	8	194.23	194.23	10	10	198.33	198.33
BB-6	3	3	177.26	178.26	8	8	194.39	194.39
BB-7	7	7	190.37	190.37	8	8	194.59	194.59
BB-8	6	6	186.61	186.61	6	6	186.61	186.61
BB-9	8	8	194.39	194.39	8	8	194.39	194.39
BB-10	6	6	186.39	186.39	9	8	195.23	194.23
BB-11	8	8	194.02	194.02	8	8	194.02	194.02
BB-12	3	3	178.08	178.08	3	3	178.08	178.08
BB-13	7	7	190.57	190.57	8	8	194.54	194.54
BB-14	3	3	178.18	178.18	3	3	178.18	178.18
BB-15	8	8	194.57	194.57	8	8	194.57	194.57
BB-16	6	6	186.34	186.34	10	10	198.66	198.66
BB-17	7	7	190.32	190.32	8	8	194.57	194.57
BB-18	3	3	177.97	177.97	7	7	190.28	190.28
BB-19	6	6	186.52	186.52	8	8	194.36	194.36
BB-20	3	3	178.19	178.19	8	8	194.48	194.48
BB-21	7	7	190.37	190.37	7	7	190.37	190.37
BB-22	7	7	190.27	190.27	8	8	194.39	194.39
BB-23	6	6	186.47	186.47	8	8	194.62	194.62
BB-24	3	3	178.36	178.36	6	6	186.55	186.55
BB-25	6	6	186.44	186.44	6	6	186.44	186.44
BB-26	3	3	178.03	178.03	6	6	186.3	186.3
BB-27	8	8	194.27	194.27	8	8	194.27	194.27
BB-28	3	3	178.09	178.09	8	8	194.3	194.3
BB-29	3	3	177.92	177.92	6	6	186.31	186.31
BB-30	3	3	178.12	178.12	6	6	186.28	186.28
bb-1	3	3	178.34	178.34	8	8	194.57	194.57
bb-2	5	5	183.61	182.61	5	5	183.61	182.61
bb-3	6	6	186.14	186.14	8	8	194.48	194.48
bb-4	3	3	178.03	178.03	6	6	186.3	186.3
bb-5	7	7	190.37	190.37	7	7	190.37	190.37
bb-6	8	8	194.33	194.33	8	8	194.33	194.33
bb-7	0	0	DNW	DNW	0	0	DNW	DNW
bb-8	7	7	190.52	190.52	10	10	198.8	198.8
bb-9	0	0	DNW	DNW	0	0	DNW	DNW
bb-10	7	7	190.71	190.71	10	10	198.66	198.66
bb-11	8	8	194.45	194.45	10	10	198.46	198.46
bb-12	6	6	186.17	186.17	6	6	186.17	186.17

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
bb-13	1	1	149.94	149.94	2	2	157.38	157.38
bb-14	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
bb-15	6	6	186.33	186.33	10	10	198.49	198.49
bb-16	7	7	190.51	190.51	10	10	198.5	198.5
bb-17	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
bb-18	6	6	185.11	186.11	6	6	185.11	186.11
bb-19	0	0	DNW	DNW	0	0	DNW	DNW
bb-20	0	0	DNW	DNW	0	0	DNW	DNW
bb-21	0	0	DNW	DNW	0	0	DNW	DNW
bb-22	0	0	DNW	DNW	0	0	DNW	DNW
bb-23	0	0	DNW	DNW	0	0	DNW	DNW
bb-24	0	0	DNW	DNW	0	0	DNW	DNW
bb-25	0	0	DNW	DNW	0	0	DNW	DNW
bb-26	0	0	DNW	DNW	0	0	DNW	DNW
RB-1	6	6	186.2	186.2	8	8	194.41	194.41
RB-2	6	6	186.2	186.2	8	8	194.41	194.41
RB-3	3	3	178.13	178.13	3	3	178.13	178.13
RB-4	5	5	180.41	182.41	8	8	194.69	194.69
RB-5	8	8	194.69	194.69	8	8	194.69	194.69
RB-6	3	3	178.48	178.48	8	8	194.75	194.75
RB-7	8	8	194.75	194.75	8	8	194.75	194.75
RB-8	3	3	178.38	178.38	8	8	194.62	194.62
RB-9	0	0	DNW	DNW	0	0	DNW	DNW
RB-10	8	8	194.62	194.62	11	11	202.67	202.67
RB-11	6	6	186.51	186.51	8	8	194.57	194.57
RB-12	8	8	194.57	194.57	8	8	194.57	194.57
RB-13	7	7	190.52	190.52	11	11	202.83	202.83
RB-14	0	0	DNW	DNW	0	0	DNW	DNW
RB-15	0	0	DNW	DNW	0	0	DNW	DNW
RB-16	6	6	186.64	186.64	6	6	186.64	186.64
RB-17	10	10	198.65	198.65	10	10	198.65	198.65
RB-18	3	3	178.32	178.32	8	8	194.59	194.59
RB-19	8	8	194.6	194.6	8	8	194.6	194.6
RB-20	0	0	Multiple peaks	Multiple peaks	0	0	Multiple peaks	Multiple peaks
RB-21	3	3	178.54	178.54	3	3	178.54	178.54
RB-23	7	7	190.57	190.57	8	8	194.54	194.54
RB-24	8	8	194.75	194.75	8	8	194.75	194.75
RB-25	8	8	194.62	194.62	8	8	194.62	194.62
RB-26	6	6	186.69	186.69	8	8	194.85	194.85
RB-27	6	6	186.68	186.68	6	6	186.68	186.68
RB-28	3	3	178.44	178.44	3	3	178.44	178.44

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
RB-29	3	3	178.5	178.5	6	6	186.63	186.63
RB-30	3	3	178.37	178.37	8	8	194.65	194.65
RB-31	6	6	186.58	186.58	6	6	186.58	186.58
RB-32	6	6	186.51	186.51	8	8	194.57	194.57
RB-33	7	7	190.71	190.71	7	7	190.71	190.71
RB-34	8	8	194.9	194.9	8	8	194.9	194.9
RB-35	6	6	186.83	186.83	8	8	194.85	194.85
RB-36	8	8	194.67	194.67	11	11	202.78	202.78
RB-37	6	6	186.74	186.74	10	10	198.69	198.69
RB-38	7	7	190.61	190.61	8	8	194.65	194.65
RB-39	88	8	194.44	194.44	8	8	194.44	194.44
RB-40	3	3	178.26	178.26	8	8	194.57	194.57
RB-41	0	0	DNW	DNW	0	0	DNW	DNW
RB-42	8	8	194.26	194.26	8	8	194.26	194.26
RB-43	3	3	178.28	178.28	8	8	194.69	194.69
RB-44	3	3	178.26	178.26	8	8	194.62	194.62
RB-45	8	8	194.51	194.51	8	8	194.51	194.51
RB-46	0	0	DNW	DNW	0	0	DNW	DNW
RB-47	8	8	194.77	194.77	10	10	198.81	198.81
RB-48	6	6	186.66	186.66	6	6	186.66	186.66
RB-49	3	3	178.59	178.59	10	10	198.8	198.8
RB-50	8	8	194.57	194.57	10	10	198.64	198.64
RB-51	8	8	194.69	194.69	8	8	194.69	194.69
RB-52	6	6	186.42	186.42	8	8	194.59	194.59
RB-53	6	6	186.37	186.37	8	8	194.51	194.51
RB-54	8	8	194.59	194.59	8	8	194.59	194.59
RB-55	7	7	190.71	190.71	7	7	190.71	190.71
RB-56	6	6	186.42	186.42	10	10	198.8	198.8
RB-57	7	7	190.62	190.62	7	7	190.62	190.62
LM-1	3	3	178.32	178.32	3	3	178.32	178.32
LM-2	3	3	178.32	178.32	8	8	194.59	194.59
LM-3	0	0	DNW	DNW	0	0	DNW	DNW
LM-4	0	0	DNW	DNW	0	0	DNW	DNW
LM-5	3	3	178.38	178.38	3	3	178.38	178.38
LM-6	3	3	178.28	178.28	3	3	178.28	178.28
LM-7	3	3	178.37	178.37	3	3	178.37	178.37
LM-8	3	3	178.42	178.42	8	8	194.62	194.62
LM-9	0	0	DNW	DNW	0	0	DNW	DNW
LM-10	3	3	178.48	178.48	3	3	178.48	178.48
LM-11	3	3	178.07	178.07	3	3	178.07	178.07
LM-12	3	3	178.22	178.22	3	3	178.22	178.22
LM-13	3	3	178.42	178.42	8	8	194.62	194.62
LM-14	3	3	178.37	178.37	3	3	178.37	178.37
LM-15	3	3	178.38	178.38	3	3	178.38	178.38

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
LM-16	3	3	178.42	178.42	3	3	178.42	178.42
LM-17	0	0	DNW	DNW	0	0	DNW	DNW
LM-18	8	8	194.88	194.88	8	8	194.88	194.88
LM-19	3	3	178.56	178.56	8	8	194.56	194.56
LM-20	3	3	178.59	178.59	3	3	178.59	178.59
LM-21	0	0	DNW	DNW	0	0	DNW	DNW
LM-22	0	0	DNW	DNW	0	0	DNW	DNW
LM-23	3	3	178.44	178.44	8	8	194.75	194.75
LM-24	3	3	178.32	178.32	8	8	194.75	194.75
LM-25	3	3	178.66	178.66	3	3	178.66	178.66

Locus Asa 16								
Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
BB-1	7	6	125.43	124.43	7	6	125.43	124.43
BB-2	7	6	125.57	124.57	7	6	125.57	124.57
BB-3	7	6	125.29	124.29	7	6	125.29	124.29
BB-4	2	2	99.55	100.55	2	2	99.55	100.55
BB-5	6	6	124.81	124.81	6	6	124.81	124.81
BB-6	6	6	124.81	124.81	6	6	124.81	124.81
BB-7	7	6	125.56	124.56	7	6	125.56	124.56
BB-8	7	6	125.8	124.8	7	6	125.8	124.8
BB-9	7	6	125.71	124.71	7	6	125.71	124.71
BB-10	7	6	125.74	124.74	7	6	125.74	124.74
BB-11	5	5	122.39	122.39	7	6	126.06	124.06
BB-12	6	6	124.69	124.69	6	6	124.69	124.69
BB-13	6	6	124.61	124.61	6	6	124.61	124.61
BB-14	6	6	124.59	124.59	6	6	124.59	124.59
BB-15	7	6	125.52	124.52	7	6	125.52	124.52
BB-16	2	2	99.37	100.37	2	2	99.37	100.37
BB-17	7	6	125.41	124.41	7	6	125.41	124.41
BB-18	2	2	99.59	100.59	2	2	99.59	100.59
BB-19	7	6	125.18	124.18	7	6	125.18	124.18
BB-20	7	6	125.66	124.66	7	6	125.66	124.66
BB-21	7	6	125.39	124.39	7	6	125.39	124.39
BB-22	7	6	125.51	124.51	7	6	125.51	124.51
BB-23	7	6	125.63	124.63	7	6	125.63	124.63
BB-24	7	6	126.05	126.05	7	6	126.05	124.05
BB-25	7	6	125.9	124.9	7	6	125.9	124.9
BB-26	7	6	125.4	124.4	7	6	125.4	124.4
BB-27	7	6	125.3	124.3	7	6	125.3	124.3
BB-28	7	6	125.47	124.47	7	6	125.47	124.47
BB-29	7	6	125.5	124.5	7	6	125.5	124.5
BB-30	7	6	126.06	124.06	7	6	126.06	124.06
bb-1	7	6	125.7	124.7	7	6	125.7	124.7
bb-2	0	0	DNW	DNW	0	0	DNW	DNW
bb-3	7	6	125.44	124.44	7	6	125.44	124.44
bb-4	6	6	124.96	124.96	6	6	124.96	124.96
bb-5	7	6	125.56	124.56	7	6	125.56	124.56
bb-6	7	6	125.44	124.44	7	6	125.44	124.44
bb-7	8	8	127.38	127.38	8	8	127.38	127.38
bb-8	6	6	124.89	124.89	6	6	126.27	124.27
bb-9	6	6	124.91	124.91	6	6	124.91	124.91
bb-10	6	6	123.9	123.9	6	6	123.9	124.9
bb-11	7	6	125.21	124.21	7	6	125.21	124.21
bb-12	6	6	124.66	124.66	6	6	124.66	124.66

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
bb-13	1	2	98.08	100.08	1	2	98.08	100.08
bb-14	7	6	125.56	124.56	7	6	125.56	124.56
bb-15	2	2	100.68	100.68	2	2	100.68	100.68
bb-16	7	6	125.43	124.43	7	6	125.43	124.43
bb-17	0	0	DNW	DNW	0	0	DNW	DNW
bb-18	0	0	DNW	DNW	0	0	DNW	DNW
bb-19	0	0	DNW	DNW	0	0	DNW	DNW
bb-20	0	0	DNW	DNW	0	0	DNW	DNW
bb-21	0	0	DNW	DNW	0	0	DNW	DNW
bb-22	0	0	DNW	DNW	0	0	DNW	DNW
bb-23	0	0	DNW	DNW	0	0	DNW	DNW
bb-24	0	0	DNW	DNW	0	0	DNW	DNW
bb-25	0	0	DNW	DNW	0	0	DNW	DNW
bb-26	0	0	DNW	DNW	0	0	DNW	DNW
RB-1	7	6	125.54	124.54	7	6	125.54	124.54
RB-2	7	6	125.38	124.38	7	6	125.38	124.38
RB-3	6	6	124.71	124.71	6	6	124.71	124.71
RB-4	7	6	125.91	124.91	7	6	125.91	124.91
RB-5	7	6	125.49	124.49	7	6	125.49	124.49
RB-6	5	5	121.56	121.56	5	5	121.56	121.56
RB-7	0	0	DNW	DNW	0	0	DNW	DNW
RB-8	7	6	125.96	124.96	7	6	125.96	124.96
RB-9	7	6	125.96	124.96	7	6	125.96	124.96
RB-10	8	8	127.36	127.36	8	8	127.36	127.36
RB-11	3	3	114.32	115.32	3	3	114.32	114.32
RB-12	8	8	127.76	127.76	8	8	127.76	127.76
RB-13	7	6	125.78	124.78	7	6	125.78	124.78
RB-14	7	6	126.14	124.14	7	6	126.14	124.14
RB-15	7	6	126.03	124.03	7	6	126.03	124.03
RB-16	0	0	DNW	DNW	0	0	DNW	DNW
RB-17	7	6	125.96	124.96	7	6	125.96	124.96
RB-18	5	5	121.65	121.65	5	5	121.65	121.65
RB-19	7	6	125.72	124.72	7	6	125.72	124.72
RB-20	6	6	124.91	124.91	6	6	124.91	124.91
RB-21	7	6	125.85	124.85	7	6	125.85	124.85
RB-23	7	6	125.05	124.05	7	6	125.05	124.05
RB-24	7	6	126.9	124.9	7	6	126.9	124.9
RB-25	6	6	124.89	124.89	6	6	124.89	124.89
RB-26	6	6	124.79	124.79	6	6	124.79	124.79
RB-27	4	4	118.86	118.86	4	4	118.86	118.86
RB-28	7	6	126.23	124.23	7	6	126.23	124.23
RB-29	0	0	DNW	DNW	0	0	DNW	DNW
RB-30	7	6	125.87	124.87	7	6	125.87	124.87
RB-31	7	6	125.83	124.83	7	6	125.83	124.83

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
RB-32	8	8	127	127	8	8	127	127
RB-33	7	6	125.93	124.93	7	6	125.93	124.93
RB-34	8	8	127.58	127.58	8	8	127.58	127.58
RB-35	7	6	126.06	124.06	7	6	126.06	124.06
RB-36	7	6	125.8	124.8	7	6	125.8	124.8
RB-37	7	6	126.76	124.76	7	6	126.76	124.76
RB-38	7	6	126.11	124.11	7	6	126.11	124.11
RB-39	6	6	124.81	124.81	6	6	124.81	124.81
RB-40	7	6	125.59	124.59	7	6	125.59	124.59
RB-41	6	6	124.94	124.94	6	6	124.94	124.94
RB-42	7	6	125.31	124.31	7	6	125.31	124.31
RB-43	7	6	125.42	124.42	7	6	125.42	124.42
RB-44	7	6	125.48	124.48	7	6	125.48	124.48
RB-45	7	6	125.4	124.4	7	6	125.4	124.4
RB-46	5	6	124.88	124.88	5	6	124.88	124.88
RB-47	7	6	125.77	124.77	7	6	125.77	124.77
RB-48	7	6	125.6	124.6	7	6	125.6	124.6
RB-49	8	8	127.25	127.25	8	8	127.25	127.25
RB-50	0	0	DNW	DNW	0	0	DNW	DNW
RB-51	7	6	125.28	124.28	7	6	125.28	124.28
RB-52	7	6	125.91	124.91	7	6	125.91	124.91
RB-53	7	6	125.42	124.42	7	6	125.42	124.42
RB-54	6	6	124.66	124.66	6	6	124.66	124.66
RB-55	7	6	125.46	124.46	7	6	125.46	124.46
RB-56	7	6	125.6	124.6	7	6	125.6	124.6
RB-57	7	6	125.52	124.52	7	6	125.52	124.52
LM-1	6	6	124.38	124.38	6	6	124.38	124.38
LM-2	5	5	121.83	121.83	5	5	121.83	121.83
LM-3	7	6	125.69	124.69	7	6	125.69	124.69
LM-4	4	4	118.53	118.53	4	4	118.53	118.53
LM-5	7	6	125.47	124.47	7	6	125.47	124.47
LM-6	7	6	125.66	124.66	7	6	125.66	124.66
LM-7	7	6	125.71	124.71	7	6	125.71	124.71
LM-8	7	6	125.81	124.81	7	6	125.81	124.81
LM-9	0	0	DNW	DNW	0	0	DNW	DNW
LM-10	0	0	DNW	DNW	0	0	DNW	DNW
LM-11	7	6	125.48	124.48	7	6	125.48	124.48
LM-12	7	6	125.59	124.59	7	6	125.59	124.59
LM-13	7	6	125.71	124.71	7	6	125.71	124.71
LM-14	7	6	125.46	124.46	7	6	125.46	124.46
LM-15	2	2	99.72	100.72	2	2	99.72	100.72
LM-16	7	6	125.85	124.85	7	6	125.85	124.85
LM-17	7	6	125.58	124.58	7	6	125.58	124.58
LM-18	8	8	127.35	127.35	8	8	127.35	127.35

Sample	Unedited Allele 1	Edited Allele 1	Unedited Size 1	Edited Size 1	Unedited Allele 2	Edited Allele 2	Unedited Size 2	Edited Size 2
LM-19	7	6	125.61	124.61	7	6	125.61	124.61
LM-20	7	6	125.91	124.91	7	6	125.91	124.91
LM-21	7	6	125.83	124.83	7	6	125.83	124.83
LM-22	0	0	DNW	DNW	0	0	DNW	DNW
LM-23	0	0	DNW	DNW	0	0	DNW	DNW
LM-24	7	6	125.75	124.75	7	6	125.75	124.75
LM-25	7	6	125.56	124.56	7	6	125.56	124.56

APPENDIX B: Microsatellite allele frequencies for alewife from all sampling locations.

Locus Asa2

Allele number (BP)	BB	RB	LM
1 (79)	0.045	0.000	0.000
2 (82)	0.000	0.087	0.000
3 (85)	0.000	0.043	0.000
4 (87)	0.000	0.087	0.000
5 (97)	0.114	0.000	0.000
6 (100)	0.091	0.022	0.000
8 (103)	0.568	0.217	0.000
9 (106)	0.091	0.022	0.000
11 (109)	0.000	0.326	0.000
13 (113)	0.000	0.152	0.000
19 (125)	0.091	0.000	0.000
22 (132)	0.000	0.022	0.000
30 (148)	0.000	0.022	0.000

Locus Asa4

Allele number (BP)	BB	RB	LM
1 (103)	0.000	0.000	0.033
3 (109)	0.000	0.000	0.033
8 (118)	0.955	0.988	0.933
10 (122)	0.015	0.000	0.000
12 (125)	0.015	0.000	0.000
13 (128)	0.015	0.012	0.000

Locus Asa8

Allele number (BP)	BB	RB	LM
1 (116)	0.000	0.000	0.021
2 (120)	0.000	0.056	0.000
3 (124)	0.014	0.000	0.000
4 (128)	0.057	0.069	0.021
6 (136)	0.786	0.667	0.958
8 (136)	0.114	0.208	0.000
10 (140)	0.029	0.000	0.000

Locus Asa9

Allele number (BP)	BB	RB	LM
1 (149)	0.011	0.000	0.000
2 (157)	0.011	0.000	0.000
3 (177)	0.170	0.150	0.789
5 (182)	0.023	0.010	0.000
6 (186)	0.227	0.190	0.000
7 (190)	0.159	0.090	0.000
8 (194)	0.307	0.460	0.211
10 (198)	0.091	0.070	0.000
11 (202)	0.000	0.030	0.000

Locus Asa16

Allele number (BP)	BB	RB	LM
2 (100)	0.111	0.000	0.048
3 (115)	0.000	0.019	0.000
4 (118)	0.000	0.019	0.048
5 (121)	0.011	0.038	0.048
6 (124)	0.856	0.827	0.810
8 (127)	0.022	0.096	0.048